

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11)

EP 1 204 073 A2

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
08.05.2002 Bulletin 2002/19

(51) Int. Cl.<sup>7</sup>: G06T 7/00

(21) Application number: 01308441.3

(22) Date of filing: 03.10.2001

(84) Designated Contracting States:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE TR

Designated Extension States:  
AL LT LV MK RO SI

(30) Priority: 27.10.2000 GB 0026343  
18.05.2001 GB 0112204

(71) Applicant: CANON KABUSHIKI KAISHA  
Tokyo (JP)

(72) Inventor: Baumberg, Adam Michael, Canon Res.  
Ctr. Europe Ltd  
Bracknell, Berkshire RG12 2XH (GB)

(74) Representative:  
Beresford, Keith Denis Lewiset al  
Beresford & Co.,  
2-5 Warwick Court,  
High Holborn  
London WC1R 5DJ (GB)

### (54) Image generation method and apparatus

(57) An image processing apparatus (2) is disclosed in which input images are processed to generate texture map data for texture rendering a generated three-dimensional computer model of object(s) appearing in the images. In order to select the portions of the images utilised, confidence data is generated indicative of the extent portions of the surface of a model are visible in each of the images. The images are then combined utilising this confidence data, where image data representative of different spatial frequencies are blended in different ways utilising the confidence data.

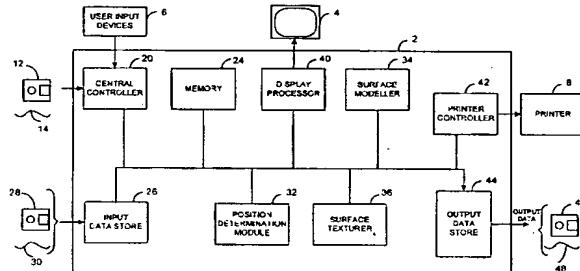


FIG. 1

EP 1 204 073 A2

## Description

[0001] The present application concerns method and apparatus for the generation of images. In particular the present application concerns method and apparatus for the generation of images of an object from a selected viewpoint from a plurality of images of the same object from other viewpoints.

5 [0002] It is known to split two-dimensional image data into a number of frequency bands and process the different frequency band images separately. When generating images of three-dimensional objects from a number of different view of the same object, it would be desirable to be able to process image data for three dimensional objects in a similar way.

[0003] One way of extending the known processing of two-dimensional images would be to process images by:

10 projecting data from the available two-dimensional images on to the surface of an object;

blurring the generated textures across the surface of the three-dimensional object to obtain a low band textured surface;

subtracting the original textured surface to obtain the corresponding high band textured surface

15 and then blending textured surfaces using the different spacial frequency information for the different bands.

[0004] However, it is not known how to perform the blurring operation for textures on the surface of a model in an efficient and consistent way. An alternative method for generating spacial frequency image data for an image of a 3D object from a selected viewpoint is therefore required.

20 [0005] In accordance with one aspect of the present invention there is provided a method of generating spacial frequency data for a specified view of an object from a plurality of images of said object comprising the steps of:

splitting available images of an object into a number of bands of spacial frequency information;

projecting the separate bands of spacial frequency information into the specified view; and generating spacial frequency data for said specified view utilising said projections.

25 [0006] By generating spacial frequency data directly from the available images of an object, a means is provided to ensure a consistent and efficient method of generating high and low frequency data for a selected image, which does not require a blurring operation for the surface of an object to be determined.

[0007] In accordance with a further aspect of the present invention there is provided an apparatus for generating spacial frequency data for a specified view of an object from a plurality of images of said object comprising:

means for splitting available images of an object into a number of bands of spacial frequency information;

means for projecting the separate bands of spacial frequency information into the specified view; and

means for generating spacial frequency data for said specified view utilising said projections.

35 [0008] Further aspects and embodiments of the present invention will become apparent with reference to the following description and drawings in which:-

40 Figure 1 is a schematic block diagram of a first embodiment of the present invention;

Figure 2 is an exemplary illustration of the position and orientation of six texture maps bounding an exemplary subject object;

Figure 3 is a schematic block diagram of the surface texturer of Figure 1;

Figure 4 is a flow diagram of the processing of the weight determination module of Figure 3;

Figure 5 is a flow diagram of processing to determine the visibility of portions of the surface of an object in input images;

Figure 6 is a flow diagram of the processing of visibility scores to ensure relatively smooth variation of the scores;

50 Figures 7 and 8 are an exemplary illustration of the processing of an area of a model ensuring smooth variation in scores;

Figure 9 is a flow diagram of the generation of weight function data;

Figures 10A and 10B are an illustrative example of a selection of triangles for associated visibility scores and generated weight function data;

55 Figure 11 is a schematic block diagram of a texture map determination module;

EP 1 204 073 A2

Figure 12 is a flow diagram of the processing of low frequency canonical projections;

Figure 13 is a flow diagram of the processing of high frequency canonical projections;

Figure 14 is a schematic block diagram of a surface texturer in accordance with a second embodiment of the present invention;

5 Figure 15 is a flow diagram of the processing of the canonical view determination module the surface texturer of Figure 14;

Figure 16 is a flow diagram of the processing of high frequency canonical projections in accordance with a third embodiment of the present invention; and

10 Figure 17 is a graph illustrating generated blend functions.

[0009] A first embodiment of will now be described in which a number of generated projected images corresponding to the same viewpoints are combined to generate composite texture map data for texture rendering a 3D computer model of object(s) appearing in the images.

15 First Embodiment

[0010] Referring to Figure 1, an embodiment of the invention comprises a processing apparatus 2, such as a personal computer, containing, in a conventional manner, one or more processors, memories, graphics cards etc., together with a display device 4, such as a conventional personal computer monitor, user input devices 6, such as a keyboard, mouse etc. and a printer 8.

20 [0011] The processing apparatus 2 is programmed to operate in accordance with programming instructions input, for example, as data stored on a data storage medium, such as disk 12, and/or as a signal 14 input to the processing apparatus 2, for example from a remote database, by transmission over a communication network (not shown) such as the Internet or by transmission through the atmosphere, and/or entered by a user via a user input device 6 such as a keyboard.

25 [0012] As will be described in more detail below, the programming instructions comprise instructions to cause the processing apparatus 2 to become configured to process input data defining a plurality of images of a subject object recorded from different view points. The input data then is processed to generate data identifying the positions and orientations at which the input images were recorded. These calculated positions and orientations and the image data are then used to generate data defining a three-dimensional computer model of the subject object.

30 [0013] When programmed by the programming instructions, processing apparatus 2 effectively becomes configured into a number of functional units for performing processing operations. Examples of such functional units and their interconnections are shown in Figure 1. The units and interconnections illustrated in Figure 1 are, however, notional and are shown for illustration purposes only to assist understanding; they do not necessarily represent the exact units and connections into which the processor, memory etc. of the processing apparatus 2 become configured.

35 [0014] Referring to the functional units shown in Figure 1, a central controller 20 processes inputs from the user input devices 6, and also provides control and processing for the other functional units. Memory 24 is provided for use by central controller 20 and the other functional units.

[0015] Data store 26 stores input data input to the processing apparatus 2 for example as data stored on a storage device, such as disk 28, as a signal 30 transmitted to the processing apparatus 2, or using a user input device 6. The input data defines a plurality of colour images of one or more objects recorded at different positions and orientations. In addition, in this embodiment, the input data also includes data defining the intrinsic parameters of the camera which recorded the images, that is, the aspect ratio, focal length, principal point (the point at which the optical axis intersects the imaging plane), first order radial distortion coefficient, and skew angle (the angle between the axes on the pixel grid; because the axes may not be exactly orthogonal).

40 [0016] The input data defining the input images may be generated for example by downloading pixel data from a digital camera which recorded the images, or by scanning photographs using a scanner (not shown). The input data defining the intrinsic camera parameters may be input by a user using a user input device 6.

45 [0017] Position determination module 32 processes the input images received by the input data store 26 to determine the relative positions and orientations of camera view points from which image data of an object represented by the image data have been obtained. In this embodiment, this is achieved in a conventional manner by identifying and matching features present in the input images and calculating relative positions of camera views utilising these matches.

50 [0018] Surface modeller 34 processes the data defining the input images and the data defining the positions and orientations at which the images were recorded to generate data defining a 3D computer wire mesh model representing the actual surface(s) of the object(s) in the images. In this embodiment this 3D model defines a plurality of triangles representing the surface of the subject object modelled.

55 [0019] Surface texturer 36 generates texture data from the input image data for rendering onto the surface model produced by surface modeller 34. In particular, in this embodiment the surface texturer 36 processes the input image data to generate six texture maps comprising six views of a subject object as viewed from a box bounding the subject object. These generated texture maps are then utilized to texture render the surface model so that images of a modelled subject object from any viewpoint may be generated. The processing of the surface texturer 36 to generate these texture maps from the input image data will be described in detail later.

[0020] Display processor 40, under the control of central controller 20, displays instructions to a user via display device 4. In addition, under the control of central controller 20, display processor 40 also displays images of the 3D computer model of the object from a user-selected viewpoint by processing the surface model data generated by surface modeller 34 and rendering texture data produced by surface

5 texturer 36 onto the surface model.

[0021] Printer controller 42, under the control of central controller 30 causes hard copies of images of the 3D computer model of the object selected and displayed on the display device 4 to be printed by the printer 8.

[0022] Output data store 44 stores the surface model and texture data therefor generated by surface modeller 34 and surface texturer 36. Central controller 20 controls the output of data from output data store 44, for example as data on a storage device, such as disk 46, or as a signal 48.

5 [0023] The structure and processing of the surface texturer 36 for generating texture data for rendering onto a surface model produced by the surface modeller 34 will now be described in detail.

#### Canonical Texture Maps

10 [0024] When a plurality of images of a subject object recorded from different viewpoints are available, this provides a large amount of data about the outward appearance of the subject object. Where images are recorded from different viewpoints, these images provide varying amounts of data for the different portions of the subject object as those portions are visible to a lesser or greater amount within the images. In order to create a model of the appearance of an object, it is necessary to process these images to generate texture data so that a consistent texture model of a subject object can be created.

15 [0025] In this embodiment this is achieved by the surface texture 36 which processes the input image data of a subject object recorded from different viewpoints in order to generate texture data for rendering the surface model produced by the surface modeller 34. In this embodiment this texture data comprises six texture maps, the six texture maps comprising views of subject object from the six faces of a cuboid centred on the subject object.

20 [0026] Figure 2 is an exemplary illustration of the position and orientation of six texture maps 50-55 bounding an exemplary subject object 56. In this embodiment the six texture maps comprise texture maps for six canonical views of an object being views of the object from the top 50, bottom 51, front 52, back 53, left 54 and right 55.

25 [0027] The six canonical views 50-55 comprise three pairs 50,51;52,53;54,55 of parallel image planes, centred on the origin of the coordinate system of the model, with each of the three pairs of image planes aligned along one of the three coordinate axes of the coordinate system respectively. The relative positions of the viewpoints of the canonical views 50-55 are then selected so that relative to the size of the model 56 of the subject object created by the surface modeller 34, the distance away from the object is selected so that the canonical views are equally spaced from the centre of the object and the extent of an object as viewed from each image plane is no more than a threshold number of pixels in extent. In this embodiment this threshold is set to be 12 pixels. Each of the texture maps for the model is then defined by a weak perspective projection of the subject object 56 onto the defined image planes.

30 [0028] After image data for each of the six canonical views 50-55 has been determined, image data of the model 56 of the subject object from any viewpoint can then be generated using conventional texture rendering techniques, where texture rendering data for each portion of the surface of the model 56 is generated utilizing selected portions of the canonical views 50-55.

35 [0029] By generating texture data for a model of a subject object in this way the total number of required texture maps is limited to the six canonical views 50-55. Furthermore, since each of the canonical views corresponds to a view of the projection of a real world object the canonical views 50-55 should be representative of realistic views of the subject object and hence be suited for compression by standard image compression algorithms such as JPEG which are optimised for compressing real world images.

40 [0030] As generated images of a model of a subject object may obtain texture rendering data from any of the six canonical texture maps, it is necessary to generate all of the texture maps in such a way that they are all consistent with one another. Thus in this way where different portions of an image of a subject object are rendered utilizing texture maps from different views no noticeable boundaries arise. As for any 3D object not all of the surface of a subject object from all six canonical views 50-55 will be visible in any single input image, it is necessary for the surface texture 36 to combine image data from the plurality of images available to generate texture maps for these six consistent views 50-55.

45 [0031] Prior to describing in detail the processing by the surface texturer 36 which enables a set of six consistent texture maps to be generated from the available input images, the structure of the surface texturer 36 in terms of notional functional processing units will now be described.

#### Structure of Surface Texturer

50 [0032] Figure 3 is a schematic block diagram of the surface texturer 36 in accordance with this embodiment of the present invention. In this embodiment, the surface texturer 36 comprises a weight determination module 58 and a texture map determination module 59.

55 [0033] In order to select portions of available image data to be utilized to generate the six canonical texture maps 50-55 the surface texture 36 utilizes position data generated by the position determination module 32 identifying the viewpoints in which data has been obtained and 3D model data output by the surface model. This position data and 3D model data is processed by the weight determination module 58 which initially determines the extent to which portions of the surface of the subject object being modelled are visible in each of the available images. The weight determination module 58 then utilizes this determination to generate weight function data identifying a relative preference for utilizing the available input images for generating texture data for each portion of the surface of the model of the subject object.

[0034] This weight function data is then passed to the texture map determination module 59 which processes the weight function data together with the position data generated by the position determination module 32, the model data generated by the surface modeller 34 and the available image data stored within the input data store 26 to generate a set of six consistent texture maps for the canonical views 50-55.

#### Generation of Weight Function Data

[0035] The processing of the weight determination module 58 generating weight function data indicative of

## EP 1 204 073 A2

relative preferences for utilizing different input images for generating texture data for different portions of a model of a subject object will now be described in detail with reference to Figures 4-9, 10A and 10B.

[0036] Figure 4 is a flow diagram of the processing of the weight determination module 58.

[0037] Initially the weight determination module 58 determines (S4-1) data indicative of the extent to which each of the triangles of the 3D model generated by the surface modeller 34 is visible in each of the images in the input store 26.

[0038] It is inevitable that for any particular input image, some portions of the subject object represented by triangles within the three dimensional model generated by the surface model 34 will not be visible. However, the same triangles may be visible in other images. Realistic texture data for each part of the surface of the 3D model can therefore only be generated by utilizing the portions of different items of image data where corresponding portions of the surface of a subject object are visible. Thus by determining which triangles are visible within each image potential sources of information for generating texture map data can be identified.

[0039] Additionally, it is also useful to determine the extent to which each triangle is visible in the available images. In this way it is possible to select as preferred sources of image data for generating portions of texture maps, images where particular triangles are clearly visible, for example in a close up images, rather than images where a triangle whilst visible is viewed only at an acute angle, or from a great distance.

[0040] Figure 5 is a flow diagram illustrating in detail of the processing of the weight determination module 58 to determine the visibility of triangles in input images, utilizing the 3D model data generated by the surface modeller 34 and the position data for the input images determined by the position determination module 32.

[0041] Initially (S5-1) the weight determination module 58 selects the first view for which position data has been generated by the position determination module 32.

[0042] The weight determination module 58 then selects (S5-2) the first triangle of the 3D model generated by the surface modeller 34.

[0043] The weight determination module 58 then (S5-3) determines a visibility value for the triangle being processed as seen from the perspective defined by the position data being processed.

[0044] In this embodiment this is achieved by utilizing conventional Open GL calls to render the 3D model generated by the surface modeller 34 as seen from the perspective defined by the position data of the image being processed. The generated image data is then utilized to calculate the visibility value.

[0045] Specifically, initially all of the triangles of the 3D model viewed from the viewpoint defined by the selected position data are rendered using a single colour with Z buffering enabled. This Z buffer data is then equivalent to a depth map. The triangle being processed is then re-rendered in a different colour with the Z buffering disabled. The selected triangle is then once again re-rendered utilizing a third colour with the Z buffering re-enabled so as to utilize the depth values already present in the Z buffer. When re-rendering the glPolygonOffset OpenGL function call is then utilized to shift the triangle slightly towards the defined camera viewpoint defined by the position data to avoid aliasing effects.

[0046] A visibility value for the triangle being processed as viewed from the viewpoint defined by the currently selected position data is then determined by calculating the number of pixels in the image rendered in the second and third colours.

[0047] Specifically, if there are any pixels corresponding to the second colour in the generated image data, these are parts of the currently selected triangle that are occluded by other triangles and hence the selected triangle is partially hidden from the viewpoint defined by the position data. A visibility value for the triangle as perceived from the defined viewpoint is then set for the triangle where:

$$\text{visibility value} = 0 \text{ if } \frac{\text{no pixels in 3rd colour}}{\text{no pixels in 2nd or 3rd colour}} < \text{threshold}$$

$$\text{visibility value} = \frac{\left[ \frac{\text{no pixels in 3rd colour}}{\text{no pixels in 2nd or 3rd colour}} \right] - \text{threshold}}{1 - \text{threshold}} \text{ otherwise}$$

where the threshold is a value in this embodiment set to 0.75.

[0048] Thus in this way, where the entirety of a triangle is visible within an image a visibility value of one is associated with the triangle and position data. Where a triangle is only slightly occluded by other triangles when viewed from the position defined by the position data being processed, i.e. the fraction of pixels rendered in the third colour relative to the total number of pixels rendered in either the second or third colour is greater than the threshold but less than one, a visibility value of less than one is associated with the triangle at image position. Where the fraction of pixels rendered in the third colour relative to the total is less than the threshold value a visibility value of zero is assigned to the triangle as viewed from the position defined by the position data.

[0049] The processing of the weight determination module 58 to colour a triangle by rendering and a triangle utilizing the stored values in the Z buffer as previously described enables triangles to be processed even if two small adjacent triangles are mapped to the same pixel in the selected viewpoint. Specifically, if triangle depths are similar for triangles corresponding to the same portion of an image, the glPolygonOffset call has the effect of rendering the second colour in the front of other triangles the same distance from the selected viewpoint. Thus when the visibility of the selected triangle is determined, the triangle is only given a lower visibility value if it is obscured by other triangles closer to the defined viewpoint and not merely obscured by the rendering of triangles at the same distance to the same portion of an image.

[0050] After the visibility value for a triangle in the selected view has been determined, the weight determination module 58 then (S5-4) calculates and stores a visibility score for the triangle in the selected

viewpoint utilizing this visibility value.

[0051] By calculating the visibility value in the manner described above, a value indicative of the extent to which a triangle in a model is or is not occluded from the selected viewpoint is determined. Other factors, however, also effect the extent to which use of particular portion image data for generating texture data may be preferred. For example, close up images are liable to contain more information about the texture of an object and therefore may be preferable sources for generating texture data. A further factor which may determine whether a particular image is preferable to use for generating texture data is the extent to which a portion of the image is viewed at oblique angle. Thus in this embodiment a visibility score for a triangle in a view is determined by the weight determination module 58 utilizing the following equation:

$$\text{visibilityscore} = \frac{\text{visibilityvalue} * \cos\theta}{\text{distance}}$$

where  $\theta$  is the angle of incidence of a ray from the optical centre of the camera defining the image plane as identified by the position data being processed to the normal of the centroid of the selected triangle and the distance is the distance between the centroid of selected triangle and the optical centre of the camera.

[0052] Thus in this way in this embodiment the amount of occlusion, the obliqueness of view and the distance between image plane and a triangle being modelled all effect the visibility score. In alternative embodiments either only some of these factors could be utilized or alternatively greater weight could be placed on any particular factor.

[0053] After a visibility score for a particular triangle as perceived from a particular viewpoint has been calculated and stored, the weight determination module 58 then (S5-5) determines whether the triangle for which data has just been stored is the last of the triangles identified by the model data generated by the surface modeller 34. If this is not the case the weight determination module 58 then proceeds to select (S5-6) the next triangle of the model data generated by the surface modeller 34 and then determines a visibility value (S5-3) and calculates and stores a visibility score (S5-4) for that next triangle.

[0054] When the weight determination module 58 determines that the triangle for which a visibility score has been stored is the last triangle of the 3D model data generated by the surface modeller 34, the weight determination module 58 then (S5-7) determines whether the currently selected position data defining the position of an input image stored within the input data store 28 is the last of the sets of position data generated by the position determination module 32. If this is not the case the weight determination module 58 then selects (S5-8) the next set of position data to generate and store visibility scores for each of the triangles as perceived from that new position (S5-2-S5-7).

[0055] Thus in this way the weight determination module 58 generates and stores visibility scores for all triangles as perceived from each of the viewpoints corresponding to viewpoints of input image data in the input data store 26.

[0056] Returning to Figure 4 after visibility scores have been stored for each of the triangles as perceived from each of the camera views corresponding to position data generated by the position determination module 32, the weight determination module 58 then (S4-2) proceeds to alter the visibility scores for the triangles perceived in each view to ensure smooth variation in the scores across neighbouring triangles.

[0057] In this embodiment of the present invention the visibility scores are utilized to generate weight functions indicative of a relative preference for using portions of different input images stored within the input data store 26 for generating texture rendering data. In order that the selection of image data from different input images does not result in the creation of noticeable boundaries where different input images are used to generate the texture data, it is necessary to generate weight function data which generates a continuous and reasonably smooth weight function so that image data from different sources is blended into one another. As in this embodiment the triangles forming the 3D model data can be of significantly different sizes, a first stage in generating such a smooth weighting function is to average weighting values across areas of the model to account for this variation in triangle size.

[0058] Figure 6 is a flow diagram of the detailed processing of the weight determination module 58 to alter the visibility score associated with triangles as perceived from a particular viewpoint. The weight determination module 58 repeats processing illustrated by Figure 6 for each of the sets of visibility score data associated with triangles corresponding to each of the views for which image data is stored within the input store 26 so that visibility score data associated with adjacent triangles as perceived from each viewpoint identified by position data determined by the position determination module 32 varies relatively smoothly across the surface of the model.

[0059] Initially (S6-1), the weight determination module 58 selects the first triangle identified by 3D model data output from the surface modeller 34.

[0060] The weight determination module 58 then (S6-2) determines whether the stored visibility score of the visibility of that triangle as perceived from the view being processed is set to zero. If this is the case no modification of the visibility score is made (S6-3). This ensures that the visibility scores associated with triangles which have been determined to be completely occluded or substantially occluded from the viewpoint being processed remain associated with a visibility score of zero, and hence are not subsequently used to generate texture map data.

[0061] If the weight determination module 58 determines (S6-2) that the visibility score of the triangle being processed is not equal to zero the weight determination module 58 then (S6-4) determines the surface area of the triangle from the 3D vertices using standard methods. The weight determination module then (S6-5) determines whether the total surface area exceeds the threshold value. In this embodiment the threshold value is set to be equal to the square of 5% of the length of the bounding box defined by the six canonical views 50-55.

[0062] If the weight determination module 58 determines that the total surface area does not exceed this threshold the weight determination module then (S6-6) selects the triangles adjacent to the triangle under consideration, being those triangles having 3D model data sharing two vertices with the triangle under consideration. The weight determination module 58 then calculates (S6-4, S6-5) the total surface for the selected triangle and those adjacent to the selected triangle to determine once again whether the number of pixels the projection of this area onto the image plane corresponding to the viewpoint against which triangles are being processed exceeds the threshold value. This process is repeated until the total surface area corresponding to the selected triangles exceeds the threshold.

## EP 1 204 073 A2

[0063] When the threshold is determined to be exceeded, the weight determination module 58 then (S6-7) sets as an initial revised visibility score for the central triangle, where the initial revised score is equal to the average visibility score associated with all the currently selected triangles weighted by surface.

[0064] A threshold value, typically set to 10% of the maximum unmodified weight is then subtracted from this revised initial visibility score. Any negative values associated with triangles are then set to zero. This threshold causes triangles associated with low weights to be associated with a zero visibility score. Errors in calculating camera position and hence visibility scores might result in low weights being assigned to triangles where surfaces are not in fact visible. The reduction of visibility scores in this way ensures that any such errors do not introduce subsequent errors in generated texture map data. The final revised visibility score for the triangle is then stored.

[0065] Figures 7 and 8 are an exemplary illustration of the processing of an area of a model in order to revise the visibility scores to ensure a smooth variation in scores regardless of the size of the triangles.

[0066] Figure 7 is an exemplary illustration of an area of a model comprising nine triangles, each triangle associated with a visibility score corresponding to the number appearing within respective triangles.

[0067] When the triangle corresponding to the central cross hatched triangle with a number 1.0 in the centre of the Figure 7 is processed initially the area of the triangle is determined. If this area is determined to be below the threshold value, the adjacent triangles, the shaded triangles in Figure 7, are selected and the area corresponding to the crossed hatched triangle and the shaded triangles is then determined.

[0068] If this value is greater than the threshold value an area weighted average to the triangles is determined after subtracting the threshold from this score and setting any negative value to zero the final value is stored as the revised visibility score for the central triangle.

[0069] Figure 8 is an exemplary illustration of the area of a model corresponding to the same area illustrated in Figure 7 after all of the visibility scores for the triangles have been processed and modified where necessary. As can be seen by comparing the values associated with triangles in Figure 7 and Figure 8, wherever a triangle is associated with a zero in Figure 7 it remains associated with a zero in Figure 8. However the variation across the remaining triangle is more gradual than the variation illustrated in Figure 7, as these new scores correspond to area weighted averages of the visibility scores in Figure 7.

[0070] Returning to Figure 6 after a revised visibility score has been determined for a triangle the weight determination module 58 then determines (S6-8) whether all of the visibility scores for the triangles in a model, as seen from a viewpoint have been processed. If this is not the case the weight determination module 58 then (S6-9) selects the next triangle for processing.

[0071] Thus in this way all of the visibility scores associated with triangles are processed so that visibility scores gradually vary across the surface of the model and drop to zero for all portions of a model which are not substantially visible from the viewpoint for which the visibility scores have been generated.

[0072] Returning to Figure 4, after the visibility scores for all of the triangles in all of the views have been amended were necessary to ensure a smooth variation of visibility scores across the entirety of the model, the weight determination module 58 then (S4-3) calculates weight function data for all of the triangles in all of the views so that a smooth weight function indicative of the visibility of the surface of the model as seen from each defined viewpoint can be generated.

[0073] The processing of the weight determination module 58 for generating weight function data for a model for a single view will now be described in detail with reference to Figures 9, 10A and 10B. This processing is then repeated for the other views so that weight function data for all of the views is generated.

[0074] Figure 9 is a flow diagram of the processing of the weight determination module 58 to generate weight function data for the surface of a model defined by 3D model data received from the surface modeller 34 for a view as defined by one set of position data output by the position determination module 32 for which visibility scores have been determined and modified.

[0075] Initially the weight determination module 58 selects (S9-1) the first triangle of the model. The weight determination module 58 then (S9-2) determines for each of the vertices of the selected triangle the minimum visibility score associated with triangles in the model sharing that vertex. These minimum visibility scores are then stored as weight function data for the points of the model corresponding to those vertices.

[0076] The weight determination module 58 then (S9-3) determines for each of the edges of the selected triangle the lesser visibility score of the currently selected triangle and the adjacent triangle which shares that common edge. These values are then stored as weight function data defining the weight to be associated with the mid point of each of the identified edges.

[0077] The weight determination module 58 then (S9-4) stores as weight function data the visibility score associated with the currently selected triangle as the weight function to be associated with the centroid of the selected triangle.

[0078] By allotting weight function data for the vertices, edges and centroids of each of the triangles for the models generated by the surface modeller 34, values for the entirety of the surface of the model of an object can then be determined by interpolating the remaining points of the surface of the model as will be described later.

[0079] Figures 10A is an illustrative example of a selection of triangles each associated with a visibility score.

[0080] Figure 10B is an example of the weight functions associated with the vertices, edges and centroid of the triangle in Figure 10A associated with a visibility score of 0.5.

[0081] As can be seen from Figure 10B the selection of weight function data for the vertices and edges of a triangle to be such that the weight function is the minimum of scores associated with triangles having a common edge or vertex ensures that the weight function associated with the edge or vertex of triangles adjacent to a triangle with a visibility score of zero are all set to zero.

[0082] By having the centre, edge and vertex of each triangle associated with weight function data, a means is provided to ensure that weight values subsequently associated with the edges of triangles can vary even when the vertices of a triangle associated with the same value. Thus for example in Figure 10B the central portion of the base of the triangle is associated with a value of 0.5 whilst the two vertices at the base of the triangle are both associated with zero. If a simpler weight function were to be utilized and weights associated with positions along the edges of a triangle were to be solely determined by weights allotted to the vertices, this variation could not occur.

[0082] Returning to Figure 9 after the weight function data has been determined and stored for the vertices, edges and centroid of the triangle currently under consideration the weight determination module 58 then determines (S9-5) whether the triangle currently under consideration is the last of the triangles in the model generated by the surface modeller 34.

[0083] If this is not the case the next triangle is then selected (S9-6) and weight function data is determined and stored for the newly selected triangle (S9-2-S9-4).

5 [0084] Thus in this way weight function data for generating a weight function that varies smoothly across the surface of a modelled object and reduces to zero for all portions of a model which are not visible from a viewpoint associated with the weight function is generated.

10 [0085] Returning to Figure 4 after weight function data has been determined for each of the camera views corresponding to position data generated by the position determination module 32 this weight function data is then (S4-4) output by the weight determination module 58 to the texture map determination module 59 so that the texture map determination module 59 can utilize the calculated weight function data to select portions of image data to be utilized to generate texture maps for a generated model.

#### Generation of Texture Maps

15 [0086] Utilizing conventional texture rendering techniques, it is possible to generate image data corresponding to projected images as perceived from each of the canonical views 50-55 where the surface of a model of a subject object is texture rendered utilizing input image data identified as having been recorded from a camera viewpoint corresponding to position data output by the position determination module 32 for that image data.

[0087] It is also possible utilizing conventional techniques to generate projected images corresponding to the projection of the surface of a model texture rendered in accordance with calculated weight functions for the surface of the model as viewed from each of the canonical views 50-55 from the output weight function data.

20 [0088] As described above, the weight function data generated output by the weight determination module 58 are calculated so as to be representative of a relative visibility of portions of the surface of a model from defined viewpoints. The projected images of the weight functions are therefore indicative of relative preferences for using the corresponding portions of projections of image data from the corresponding viewpoints to generating the texture data for the texture map for each of the canonical views. These projected weight function images, hereinafter referred to as canonical confidence images can therefore be utilized to select portions of projected image data to blend to generate output texture maps for canonical views 50-55.

25 [0089] Figure 11 is a schematic diagram of notional functional modules of the texture map determination module 59 for generating canonical texture maps for each of the canonical views 50-55 from image data stored within the input data store 26, position data output by the position determination module 32, 3D model data generated by the surface modeller 34 and the weight function data output by the weight determination module 58.

[0090] The applicants have appreciated that generation of realistic texture for a model of a subject object can be achieved by blending images in different ways to average global lighting effects whilst maintaining within the images high frequency details such as highlights and shadows.

30 [0091] Thus in accordance with this embodiment of the present invention the texture generation module 59 comprises a low frequency image generation module 60 for extracting low frequency image information from image data stored within the input data store 26, an image projection module 62 for generating high and low frequency canonical image projections; a confidence image generation module 64 for generating canonical confidence images; a weighted average filter 66 and a blending module 68 for processing high and low frequency canonical projections and the canonical confidence images to generate blended high and low frequency canonical images; and a re-combination of output module 70 for combining the high and low frequency images and outputting the combined images as texture maps for each of the canonical views 50-55.

[0092] In order to generate high and low frequency canonical projections of each of the input images stored within the input data store 26 projected into each of the six canonical views, initially each item of image data in the input data store 26 is passed to the low frequency image generation module 60.

35 [0093] This module 60 then generates a set of low frequency images by processing the image data for each of the views in a conventional way by blurring and sub-sampling each image. In this embodiment the blurring operation is achieved by performing a Gaussian blur operation in which the blur radius is selected to be the size of the projection of a cube placed at the centre of the object bounding box defined by the six canonical views 50-55 and whose sides are 5% of the length of the diagonal of the bounding box. The selection of the blur radius in this way ensures that the radius is independent of image resolution and varies depending upon whether the image is a close up of a subject object (large radius) or the subject appears small in the image (small radius).

40 [0094] These low frequency images are then passed to the image projection module 62 together with copies of the original image data stored within the input data store 26, position data for each of the images as determined by the position determination module 32 and 3D model data for the model generated by the surface modeller 34.

[0095] For each of the input images from the input data store 26 the image projection module then utilizes the position data associated with the image and the 3D model data output by the surface modeller 34 to determine calculated projections of the input images as perceived from the six canonical views 50-55 by utilizing standard texture rendering techniques.

45 [0096] In the same way the image projection module 62 generates for each of low frequency images corresponding to image data processed by the low frequency image generation module 60, six low frequency canonical projections for the six canonical views 50-55.

[0097] Six high frequency canonical image projections for each image are then determined by the image projection module 62 by performing a difference operation, subtracting the low frequency canonical image projections for an image as viewed from a specified canonical view from the corresponding image projection of the raw image data for that view.

50 [0098] The low frequency canonical projections and high frequency canonical projections of each image for each of the six canonical views are then passed to the weighted average filter 66 and the blending module 68 for processing as will be detailed later.

[0099] By processing the image data from the input data store 26 in this way it is possible to ensure that the

## EP 1 204 073 A2

processing of images generates high and low frequency canonical projections of each image that are consistent with one another. In contrast, if a blurring operation is performed upon canonical image projections created from image data stored within the input data store 26, as the pixel data in these generated projections is normally dependent upon different regions of the original image data, the blurring operation will not be consistent across all six canonical images and hence will introduce errors into the generated texture maps.

[0100] The confidence image generation module 64 is arranged to receive 3D model data from the surface modeller 34 and weight function data from the weight function determination module 58. The confidence image generation module 64 then processes the 3D model data and weight function data to generate for each of the weight functions associated with viewpoints corresponding to the viewpoints of each input images in the input data store 26, a set of six canonical confidence images for the six canonical views 50-55.

[0101] Specifically each triangle for which weight function data has been stored is first processed by linking the midpoint of each edge of the triangle to the midpoints of the two other edges and linking each of the midpoints of each edge and the centroid of the triangle. Each of these defined small triangles, is then projected into the six canonical views with values for points corresponding to each part of the projection of each small triangle being interpolated using a standard OpenGL "smooth shading" interpolation from the weight function data associated with the vertices of these small triangles.

[0102] Thus in this way for each of the six canonical views, a canonical confidence image for each weight function is generated. The pixel values of canonical confidence images generated in this way are each representative of the relative preference for utilizing corresponding portions of canonical projections of image data representative of the viewpoint for which the weight function was generated to generate texture map data. These canonical confidence images are then passed to the weighted average filter 66 and the blending module 68, so that texture data for the six canonical views can be generated utilizing the data identifying preferred sources for generating portions of texture map data.

[0103] Thus after the processing of image data and weight function data by the low frequency image generation module 60, the image projection module 62 and the confidence image generation module 64, the weighted average filter 66 and blending module 68 receive, for each item of image data stored within the input data store 26 a set of six canonical confidence images identifying the extent to which portions of projected images derived from a specified image are to be preferred to generate texture data and a set of six projections of either the high or low frequency images corresponding to the item of image data. The weighted average filter 66 and the blending module 68 then proceed to process the confidence images and associated low and high frequency canonical projections in turn for each of the canonical views as will now be described in detail.

### Processing of Low Frequency Canonical Projections

[0104] Figure 12 is a flow diagram of the processing of low frequency canonical projections and associated canonical confidence images for a specified canonical view for each of a set of input images stored within the input data store 26.

[0105] Initially (S12-1) the weighted average filter 66 selects a first low frequency canonical projection and its associated confidence image. This is then made the basis of an initial low frequency canonical image to be generated by the weighted average filter 66. The weighted average filter 66 then selects (S12-2) the next low frequency projection for the same canonical view together with the projection's associated confidence image.

[0106] The weighted average filter 66 then (S12-3) determines as a new low frequency canonical image a canonical image comprising for each pixel in the image a weighted average of the current low frequency canonical image and the selected low frequency canonical projection weighted by the confidence scores utilising the following formula:

$$\text{pixel value} = \frac{C_i}{C + C_i} * \left[ \begin{array}{l} \text{Pixel Value in} \\ \text{selected projection} \end{array} \right] + \frac{C}{C + C_i} * \left[ \begin{array}{l} \text{Pixel Value} \\ \text{in canonical image} \end{array} \right]$$

40 where C is the current confidence score associated with the pixel being processed for the canonical image and  $C_i$  is the confidence score associated with the pixel in the confidence image associated with the selected projection.

[0107] The confidence score for the pixel in the canonical image is then updated by adding the confidence score for the latest projection to be processed to the current confidence score. That is to say the new confidence score  $C_{\text{new}}$  is calculated by the equation

$$C_{\text{new}} = C_{\text{old}} + C_i$$

45 where  $C_{\text{old}}$  is the previous confidence score associated with the current pixel and  $C_i$  is the confidence score of the current pixel in the projection being processed.

[0108] The weighted average filter 66 then (S12-4) determines whether the latest selected low frequency canonical projection is the last of the low frequency canonical projections for the canonical view currently being calculated. If this is not the case the weighted average filter 66 then proceeds to utilize the determined combined image to generate a new combined image utilizing the next confidence image and associated low frequency image projection (S12-2-S12-4).

[0109] When the weighted average filter 66 determines (S12-4) that the last of the low frequency canonical projections for a specified canonical view 50-55 has been processed the weighted average filter 66 outputs (S12-5) as a blended canonical low frequency image the image generated utilizing the weighted average of the last projected image processed by the average weighted filter 66.

55 [0110] Thus in this way the weighted average filter 66 enables for each of the canonical views 50-55 a low frequency image to be created combining each of the low frequency canonical projections weighted by confidence scores indicative of the portions of the images identified as being most representative of the surface texture of a

model in the canonical confidence images. As the low frequency images are representative of average local colour of surfaces as effected by global lighting effects, this processing by the weighted average filter 66 enables canonical low frequency images to be generated in which these global lighting effects are averaged across the best available images where greater weight is placed on images in which portions of a model are most easily viewed giving the resultant canonical low frequency images a realistic appearance and a neutral tone.

5 Processing of High Frequency Canonical Projections

[0111] Figure 13 is a flow diagram of the processing of canonical confidence images and associated high frequency canonical projections for generating a canonical high frequency image for one of the canonical views 50-55. This processing is repeated for each of the canonical views 50-55.

[0112] The processing of the blending module 68 for generating high frequency canonical images is identical to that of the processing of the weighted average filter 66 except instead of determining weighted averages of pixels in the canonical projections and canonical images, in this embodiment the blending module 68 selects as canonical image data pixel data, pixel data associated with the greatest confidence scores.

[0113] Specifically, after selecting an initial high frequency projection and associated confidence image to be the initial high frequency canonical image and associated confidence image (S13-1) and selecting the next projection and confidence image for processing (S13-2), the blending module 68 then updates the pixels in the current high frequency canonical image (S13-3).

[0114] This is achieved by the blending module 68 selecting for each pixel in the image either pixel data for the corresponding pixel in the high frequency canonical projection being processed or the pixel data for the current high frequency canonical image. The selection is made for each pixel by determining whether the corresponding pixel in the confidence image associated with the high frequency projection being processed has a greater confidence score than the confidence score associated with the pixel in the current canonical image. The pixel data associated with the highest score is then utilised as pixel data for the canonical image. The confidence score associated with the pixel in the updated canonical image is then set to the greater of the confidence scores for the pixel as identified by either the confidence image associated with the canonical image or the projection currently be processed.

[0115] The blending module 68 then determines (S13-4) whether all of the high frequency image projections for a canonical user have been processed. If this is not the case then the next projection and confidence image are utilised to update the high frequency canonical image being generated (S13-2-S13-4). When all of the canonical projections have been processed the final canonical image is then output (S13-5).

[0116] Thus in this way a high frequency canonical image is generated utilising all of the available high frequency image projections where each pixel of the high frequency canonical image corresponds to pixel data from the available high frequency image projections associated with the greatest of the confidence scores for that pixel. As the confidence scores are indicative of the relative goodness for selecting pixels for generating image data, the high frequency canonical images therefore correspond to a patchwork of the best available pixel data for each part of the high frequency canonical images.

[0117] Thus in this way six low frequency canonical images for the six canonical views 50-55 and six high frequency canonical images are generated by the weighted average filter 66 and the blending module 68 respectively. In the canonical low frequency images the global lighting effects are averaged across the available low frequency canonical projections in proportion to the relative preference for utilizing portions of an image as identified by the canonical confidence images associated with the projections. In the high frequency canonical images, high frequency canonical projections are selected in the above described manner to ensure that high frequency detail and contrast in the canonical images is maintained when the different high frequency canonical projections are combined.

[0118] When the six canonical low frequency images and six canonical high frequency images are then received by the recombination output module 70 for each canonical view 50-55 the recombination and output module 70 generates a single canonical texture map by adding the high and low frequency images for each view. The combined canonical texture maps are then output by the recombination and output module 70 for storage in the output data store 44 together with the 3D model data generated by the surface module 34.

[0119] The output canonical texture maps can then be utilized to texture render image representations of the model generated by the surface modeller 34 by associating each of the triangles identified by the three-dimensional model data with one of the texture maps. In this embodiment the selection of texture data to texture render each triangle is determined by selecting as the map to be utilized the map in which the triangle has the greatest visibility score in which the visibility score is calculated for each canonical view 50-55 in the same way as has previously been described in relation to Figure 5. Texture co-ordinates for texture rendering the model generated by the surface modeller 34 are then implicitly defined by the projection of the triangle onto the selected canonical view 50-55.

[0120] Images of the model of the subject object can then be obtained for any viewpoint utilizing the output canonical texture maps and the 3D model data stored in the output data store 44, for display on the display device 4 and a hard copies of the generated images can be made utilizing the printer. The model data and texture map stored in the output data store 44 can also be output as data onto a storage device 46,48.

50 Second Embodiment

[0121] In the first embodiment of the present invention, an image processing apparatus was described in which input data defining a plurality of images of a subject object recorded from different viewpoints were processed to generate texture data comprising six canonical texture maps being views for a cuboid bounding a model of the subject object. Although generating texture maps in this way ensures that the total number of required view is limited to six, this does not guarantee that every triangle within the model is visible in at least one of the canonical views. As a result, some portions of a model may be texture rendered utilizing surface information corresponding to a different portion of the surface of the model. In this embodiment of the present invention an image processing apparatus is described in which texture data is generated to ensure that all triangles within a 3D model

EP 1 204 073 A2

are rendered utilizing texture data for the corresponding portion of the surface of a subject object, if image data representative of the corresponding portion of the subject object is available.

[0122] The image processing apparatus in accordance with this embodiment of the present invention is identical to that of the previous embodiment except the surface textruer 36 is replaced with a modified surface textruer 80.

[0123] Figure 14 is a schematic block diagram of the notional functional modules of the modified surface textruer module 80. The modified surface textruer 80 comprises a canonical view determination module 82; a weight determination module 58 identical to the weight determination module 58 of the surface textruer 36 of the previous embodiment and a modified texture map determination module 86.

[0124] In this embodiment the canonical view determination module 82 is arranged to determine from the 3D model data output by the surface modeller 34 whether any portions of the surface of the 3D model are not visible from the six initial canonical views. If this is determined to be the case the canonical view determination module 82 proceeds to define further sets of canonical views for storing texture data for these triangles. These view definitions are then passed to the texture map determination module 86. The modified texture map determination module 86 then generates canonical texture map data for the initial canonical views and these additional canonical views which are then utilized to texture render the surface of the model generated by the surface modeller 34.

[0125] The processing of the canonical view determination module 82 in accordance with this embodiment of the present invention will now be described with reference to Figure 15 which is a flow diagram of the processing of the canonical view determination module 82.

[0126] Initially (S15-1) the canonical view determination module 82 selects a first triangle of the model generated by the surface modeller 34. The canonical view determination module 82 then (S15-2) generates a visibility value and visibility score for the triangle as perceived from each of the six canonical views in a similar manner to which the visibility value is generated by a weight determination module 58 as has previously been described in relation to the previous embodiment.

[0127] The canonical view determination module 82 then (S15-3) determines whether the visibility value associated with the triangle being processed is greater than a threshold value indicating that the triangle is at least substantially visible in at least one of the views. In this embodiment this threshold value is set to zero so that all triangles which are at least 75% visible in one of the views are identified as being substantially visible.

[0128] If the canonical view determination module 82 determines that the triangle being processed is visible i.e. has a visibility value of greater than zero in at least one view, the canonical view determination module selects from the available canonical views the view in which the greatest visibility score as generated in the manner previously described is associated with the triangle. Data identifying the canonical view in which the selected triangle is most visible is then stored. Thus in this way data identifying the canonical view in which texture data for rendering the triangle is subsequently to be utilized is generated.

[0129] The canonical view determination module 82 then (S15-5) determines whether the currently selected triangle corresponds to the last triangle of the 3D model generated by the surface modeller 34. If this is not the case the canonical view determination module 82 then proceeds to select the next triangle and determine visibility values and visibility scores for that next triangle (S15-6-S15-5).

[0130] When the canonical view determination module 82 determines that all of the triangles have been processed (S15-5) the canonical view determination module 82 then (S15-7) determines whether data identifying a best view has been stored for all of the triangles.

[0131] If this is not the case the canonical view determination module 82 generates and stores a list of triangles for which no best view data has been stored (S15-8) the canonical view determination module 82 then proceeds to establish the visibility of these remaining triangles in the six canonical views in the absence of the triangles for which best view data has already been stored (S15-1-S15-7). This process is then repeated until best view data is stored for all of the triangles.

[0132] When the canonical view determination module 82 determines (S15-7) that best view data has been stored for all the triangles the canonical view determination module 82 then outputs (S15-9) view definition data comprising for each triangle the canonical view in which it is most visible together with the generated lists of the triangles not visible canonical views. These lists and identified views in which the triangles are best represented are then output and passed to the modified texture map determination module 86.

[0133] When the modified texture map determination module 86 processes the 3D model data, weight function data, image data and position data this generates the six original canonical views in the same manner as has been described in relation to the first embodiment.

[0134] However, additionally the texture map determination module 86 generates for the triangles not visible in the original six canonical views as identified by the lists output by the canonical view determination module 82 further texture maps corresponding to the projection of the triangles identified by the lists of triangles generated by the canonical view determination module 82 in the six canonical views in the absence of the other triangles. These additional canonical views are generated from the high and low frequency canonical projections of image data generated by the image projection module 62 of the modified texture map determination module 86 and additional canonical confidence images being canonical projections of only the triangles identified by the lists of triangles output by the canonical view determination module 82. These additional confidence images being calculated for the triangles identified in the lists utilizing the weight function data generated by the weight determination module 58 in the same way as has previously been described in relation to the first embodiment.

[0135] Thus in this way additional texture maps are generated for those portions of the surface of the model which are not apparent in the original six canonical views. As these partial additional texture maps are generated in a similar way to the main canonical views they should also be representative of projections of portions of an image and hence suitable for compression utilizing the standard image compression techniques.

[0136] The original canonical views together with these additional canonical views can then be utilized to texture render the model where the texture maps utilized to render each triangle is selected utilizing the lists and the data identifying the best views as are output by the canonical view determination module 82.

55 Third Embodiment

[0137] A third embodiment of the present invention will now be described. In the first embodiment a blending

EP 1 204 073 A2

module 68 was described which selected pixel data corresponding to the pixels associated with the highest confidence scores for the corresponding pixel in associated confidence images. A problem with such a selection system is that boundaries can develop in the generated texture map data indicative of the selection of pixel data from different image sources. The obvious solution to this problem performing an average operation such as that performed by the weighted average filter 66 is undesirable as this results in loss of contrast within the texture maps data. In this embodiment the blending module 68 is arranged to process high frequency canonical projections in a way which avoids noticeable boundaries arising within texture map data, whilst maintaining image contrast.

[0138] Figure 17 is a flow diagram of the processing canonical confidence images and associated high frequency canonical image projections by the blending module 68 in accordance with this embodiment of the present invention.

[0139] Initially (S16-1) the blending module 68 selects a first one of the high frequency canonical projections for the canonical view for which a canonical high frequency image is to be generated and sets as an initial canonical high frequency image an image corresponding to the selected first high frequency projection.

[0140] The blending module 68 then (S16-2) selects the next high frequency canonical projection for that canonical view and its associated canonical confidence image for processing.

[0141] The first pixel within the canonical high frequency image is then selected (S16-3) and the blending module 68 then (S16-4) determines for the selected pixel in the high frequency canonical image a difference value between the pixel in the current high frequency canonical image and in the corresponding pixel in the high frequency canonical projection being processed.

[0142] In this embodiment, where the high frequency canonical projections and canonical high frequency image comprise colour data, this difference value may be calculated by determining the sum of the differences between corresponding values for each of the red, green and blue channels for the corresponding pixels in the selected high frequency canonical projection and the canonical high frequency image being generated.

[0143] Alternatively, where the image being processed comprises grey scale data, this difference value could be determined by calculating the difference between the pixel in the canonical high frequency image and the corresponding pixel in the high frequency canonical projection being processed.

[0144] It will be appreciated that more generally in further embodiments of the invention where colour data is being processed any suitable function of the red, green and blue channel values for corresponding pixels in the high frequency canonical projection being processed and the canonical high frequency image being generated could be utilized to generate a difference value for a pixel.

[0145] After the difference value for the pixel being processed has been determined, the blending module 68 then (S16-5) determines a blend function for the selected pixel from the confidence scores associated with the pixel and the determined difference value for the pixel.

[0146] In this embodiment this blend function is initially determined by selecting a gradient G in dependence upon the determined difference value for the pixel. Specifically, in this embodiment the value G is calculated by setting

$$G = \frac{D}{D_0} \text{ for } D \geq D_0$$

$$G = 1 \text{ for } D < D_0$$

30 where D is the determined difference value for the pixel and  $D_0$  is a blend constant fixing the extent to which weighting factors are varied due to detected differences between images. In this embodiment where the difference values D are calculated from colour image data for three colour channels each varying from 0 to 255, the value of  $D_0$  is set to be 60.

[0147] An initial weighting fraction for calculating a weighted average between the pixel being processed in the current canonical high frequency image and the corresponding pixel in the high frequency canonical projection being processed is then set by calculating:

$$W_0 = \left( \frac{C_i}{C_i + C_j} - 0.5 \right) G + 0.5$$

40 where  $C_i(C_i + C_j)$  is the relative confidence score associated with the pixel being processed in the confidence image associated with the selected high frequency canonical projection being the ratio of the confidence score  $C_i$  associated with the pixel in the current image being processed to the confidence score associated with the pixel by the canonical image and G is the gradient determined utilizing the determined difference value D for the pixel.

[0148] A final weighting fraction is then determined by setting the weighting value W with:

$$W = 0 \text{ if } W_0 < 0$$

$$W = 1 \text{ if } W_0 > 1; \text{ and}$$

$$W = W_0 \text{ if } 0 \leq W_0 \leq 1$$

[0149] Figure 14 is a graph of a generated blend function illustrating the manner in which the final weighting value W calculated in this way varies in dependence upon the relative confidence score  $C_i/(C_i + C_j)$  for a pixel being processed and the determined difference value D for the pixel.

[0150] In the graph as illustrated, solid line 72 indicates a graph for calculating W for weighting values where D is equal to a high value, for example in this embodiment where D ranges between 0 and 765 a value of over 700.

[0151] As such the graph indicates that where the relative confidence score for a particular source is low a weighting value of zero is selected and for high pixel confidence scores, a weighting factor of one is selected. For intermediate pixel confidence scores an intermediate weighting factor is determined, the weighting factor increasing as the confidence score increases.

[0152] The dotted line 74 in Figure 17 is an illustration of the blend function where the calculated value D

EP 1 204 073 A2

for the difference between the image data of the first and second images for a pixel is less than or equal to the blend constant  $D_0$ . In this case the blend function comprises a function which sets a weighting value equal to the pixel relative confidence score for the pixel.

[0153] For intermediate values of D, the blend function varies between the solid line 72 and the dotted line 74 with the thresholds below which or above which a weighting value of zero or one is output reducing and increasing as D decreases. Thus as D decreases the proportion of relative confidence scores for which a weighting value other than zero or one is output increases.

[0154] Returning to Figure 16, after a final weighting value W has been determined the blending module 68 then proceeds to calculate and store blended pixel data (S16-6) for the pixel under consideration. In this embodiment, where the image data comprises colour image data, the calculated blended pixel data is determined for each of the three colour channels by:

10

$$\text{New pixel Value} = W \left[ \begin{array}{l} \text{pixel value in} \\ \text{selected image} \end{array} \right] + (1 - W) * \left[ \begin{array}{l} \text{pixel value in} \\ \text{canonical image} \end{array} \right]$$

15

[0155] The confidence value associated with the blended pixel is then also updated by selecting and storing as a new confidence score for the pixel the greater of the current confidence score associated with the pixel in the canonical image and the confidence score of the pixel in the projection being processed.

[0156] The effect of calculating the pixel data for pixels in the high frequency canonical image in this way is to make the pixel data dependent upon both the difference data and the confidence data associated with the processed pixel in the selected high frequency canonical projection. Specifically, where the difference data for a pixel is low (i.e. less than or equal to the blend constant  $D_0$ ) calculated pixel data corresponding to a weighted average proportional to the relative confidence score for the pixel in the associated canonical image is utilized. Where difference data is higher (i.e. greater than the blend constant  $D_0$ ) a pair of threshold values are set based upon the actual value of the difference data. Pixel data is then generated in two different ways depending upon the relative confidence score for the pixel in the associated canonical image. If the relative confidence score is above or below these threshold values either only the original pixel data for the canonical image or the pixel data for the selected high frequency canonical projection is utilized as a composite image data. If the relative confidence score is between the threshold values a weighted average of the original pixel data and pixel data from the selected projection is utilized.

[0157] After the generated pixel data has been stored, the blending module 68 then determines (S16-7) whether the pixel currently under consideration is the last of the pixels in the canonical image. If this not the case the non-linear averaging filter then (S16-8) selects the next pixel and repeats the determination of a difference value (S16-4) a blend function (S16-5) and the calculation and storage of pixel data (S16-6) for the next pixel.

[0158] If after generated pixel data and a revised confidence score has been stored for a pixel the blending module 68 determines (S16-7) that blended pixel data has been stored for all of the pixels of the high frequency canonical image, the blending module 68 then determines (S16-9) whether the selected high frequency canonical projection and associated confidence image is the last high frequency canonical projection and confidence image to be processed. If this is not the case the next high frequency canonical projection is selected (S16-2) and the canonical high frequency image is updated utilizing this newly selected projection (S16-3-S16), and its associated confidence image.

[0159] When the blending module 68 determines that the latest high frequency canonical projection utilized to update a canonical high frequency image is the last of the high frequency canonical projections, the blending module 68 then (S16-10) outputs as a high frequency image for the canonical view 50-55 the high frequency canonical image updated by the final high frequency projection.

[0160] The applicants have realized that the undesirable effects of generating composite image data by determining a weighted average of pixels from high frequency image data including data representative of details arise in areas of images where different items of source data differ significantly. Thus for example, where highlights occur in one source image at one point but not at the same point in another source image, averaging across the images results in a loss of contrast for the highlight. If the highlights in the two images occur at different points within the two source images, averaging can also result in the generation of a 'ghost' highlight in another portion of the composite image. Similarly, where different items of source data differ significantly other undesirable effects occur at points of shadow, where an averaging operation tends to cause an image to appear uniformly lit and therefore appear flat.

[0161] In contrast to areas of difference between source images, where corresponding points within different source images have a similar appearance, an averaging operation does not degrade apparent image quality. Performing an averaging across such areas of composite image, therefore provides a means by which the relative weight attached to different sources of image data can be varied so that the boundaries between areas of a composite image obtained from different source images can be made less distinct.

[0162] Thus in accordance with this embodiment of the present invention weighting factors are selected on the basis of both confidence data and difference data for a pixel so that, relative to weighting factors proportional to confidence data only as occurs in the weighted average filter 66, in the blending module 68, where determined differences between the source images are higher, the weighting given to preferred source data is increased and the corresponding weighting given to less preferred data is decreased.

[0163] Thus, where both confidence data and difference data are high, a greater weight is given to a particular source or alternatively only the preferred source data is utilized to generate composite image data. As these areas correspond to areas of detail and highlights or shadow in the high frequency images the detail and contrast in the

resultant generated data is maintained.

[0164] In areas where there is less difference between different sources of image data, weighting values proportional to confidence data are not modified at all or only slightly so that the boundaries between portions of image data generated from different images are minimised across those areas. Thus in this way composite image data can be obtained which maintains the level of contrast and detail in original high frequency source data, whilst minimising apparent boundaries between portions of composite image generated from different items of source data.

5 [0165] In the above embodiments, a set of images of a subject object is described as being utilized to generate 3D model data and texture render data for generating representations of that subject object. However, texture render data obtained utilizing images of one object could for example be utilized to generate representations of a second object. The mixing of a model obtained for one object and texture derived from a second object could be utilized to establish how the application of texture from different objects varies the appearance of a model.

10 [0166] Specifically, confidence images of a first object from any number of views can be generated in a similar manner as has previously been described. Source images of a different object can then be processed using the confidence images as if the source images had been generated. The processing of the source images in this way would then generate texture data in which the texture in the source images was blended smoothly across the surface of the first object.

15 [0167] Thus for example a number of confidence images could be generated for views of a model. The confidence images of the model viewed from viewpoints to one side of the model might be associated with an image of a first texture for example a wood grain. The remaining confidence images of the model might be associated with a second texture e.g. marble. Processing these two source images with the generated confidence images would generate texture data in which one side of the model appeared having a wood grain texture and the other a marble texture with the two textures being smoothly blended across the surface of the model.

20 [0168] Alternatively, instead of using only a limited number of texture images, images of a different object as viewed from viewpoints for which confidence images were generated could be used. Thus for example, images of a model first object could be generated as has previously been described for a number of viewpoints. Confidence images for a second object could be generated for those corresponding viewpoints and texture data for the second object obtained by processing the images of the first object utilizing the confidence images of the second object.

25 [0169] Although the embodiments of the invention described with reference to the drawings comprise computer apparatus and processes performed in computer apparatus, the invention also extends to computer programs, particularly computer programs on or in a carrier, adapted for putting the invention into practice. The program may be in the form of source or object code or in any other form suitable for use in the implementation of the processes according to the invention. The carrier be any entity or device capable of carrying the program.

30 [0170] For example, the carrier may comprise a storage medium, such as a ROM, for example a CD ROM or a semiconductor ROM, or a magnetic recording medium, for example a floppy disc or hard disk. Further, the carrier may be a transmissible carrier such as an electrical or optical signal which may be conveyed via electrical or optical cable or by radio or other means.

[0171] When a program is embodied in a signal which may be conveyed directly by a cable or other device or means, the carrier may be constituted by such cable or other device or means.

35 [0172] Alternatively, the carrier may be an integrated circuit in which the program is embedded, the integrated circuit being adapted for performing, or for use in the performance of, the relevant processes.

### Claims

1. A method of generating frequency component image data representative of image information of a plurality of different spacial frequencies of an image of at least part of a subject from a predefined view, the method comprising steps of:

35 receiving a plurality of images comprising image data representative of a subject from a plurality of view points;

40 for each of said plurality of received images, determining frequency component image data representative of image information of a plurality of different spacial frequencies within said received images and determining projections of said frequency component image data projected to said predefined view; and

45 for each of said plurality of different spacial frequencies, utilising said projections of frequency component image data to generate frequency component image data representative of spacial frequency information of an image of at least part of said subject from said predefined view.

50 2. A method in accordance with claim 1, wherein the determination of frequency component image data representative of a lower spacial frequency is determined by performing a blurring and sub-sampling operation utilizing said received image data.

55 3. A method in accordance with claim 2 wherein the determination of frequency component image data representative of a higher frequency is determined by performing a differencing operation utilising said lower frequency component image data and said received image data.

4. A method in accordance with any preceding claim wherein said projections of said frequency component image data are determined by performing the steps of:

55 determining from said plurality of received images the relative positions of the viewpoints of the subject represented by said images;

determining from said plurality of images and said determined positions, model data representative of the surface of said subject; and

determining said projections utilising said position data and said model data of the surface of said subject.

5. A method in accordance with claim 4, wherein said determination of position data indicative of the relative position of said plurality of viewpoints for said received images comprises matching features within said images to determine relative position and orientation of the viewpoints.
6. A method in accordance with claims 4 or 5, wherein said generation of model data comprises for each of said plurality of received images the steps of:
  - determining the portions of an image corresponding to the subject; and
  - 10 generating model data utilising a determination of the perimeters of said portion of said image and position data indicative of the viewpoint for said image.
7. A method in accordance with any preceding claim wherein said generation of frequency component image data comprises for each spatial frequency for which frequency component image data is to be generated the step of:
  - 15 utilising each portion of said projections of frequency component image data to determine frequency component image data representative of spatial frequency information of said spatial frequency for corresponding portions of said image.
8. A method in accordance with claim 7, wherein said generation of frequency component image data for different spatial frequencies for an image are generated from said corresponding portions of projections of frequency component image data in different ways for different spatial frequencies.
- 20 9. A method in accordance with claims 7 and 8, wherein said generation of frequency component image data further comprises the steps of:
  - 25 for each of said plurality of images, determining visibility data indicative the relative visibility of portions of the surface of said subject from said viewpoints;
  - associating portions of said projections of frequency component image data, with said visibility data; and
  - utilising said visibility data associated with portions of said projections to determine frequency component image data representative of frequency image information of an image of said subject from said predefined views.
- 30 10. A method in accordance with claim 9, wherein said generation of frequency component image data for at least one of said different spatial frequencies is generated by calculating for each portion of said image data for said spatial frequency a weighted average of corresponding portions of projections of frequency component image data for said portion of said image, weighted by said visibility data associated with said portions of image data.
- 35 11. A method in accordance with claim 9, wherein said generation of frequency component image data for at least one of said different spatial frequencies is generated by selecting for each pixel of said image data for said spatial frequency said corresponding portions of projections of frequency component image data associated with visibility data indicative of the highest visibility.
- 40 12. A method of generating image data representative of an image of a subject from a predefined viewpoint, the method of comprising the steps of:
  - 45 generating frequency component image data representative of image information of a plurality of different spatial frequencies of an image of a subject from said predefined view in accordance with a method of any preceding claim; and
  - generating image data representative said image of said subject utilising said frequency component image data.
- 50 13. A method in accordance with claim 12, wherein said generation of image data comprises for each portion of said image determining the sum of frequency component image data for said different spatial frequencies.
14. A method in accordance with claims 12 and 13 further comprising the step of outputting said generated image data.
- 55 15. A method of generating texture map data for texture rendering a model of a subject comprising the steps of:
  - generating a plurality of images of a subject from a plurality of predefined views in accordance with the method of claim 12 or claim 13.

EP 1 204 073 A2

16. A method in accordance with claim 15, wherein said plurality of predefined views comprise views of a subject from viewpoints indicative of views from the surface of a cuboid bounding said subject.

17. A method of generating image data of an image of a subject from a selected viewpoint, the method comprising the steps of:

5 generating texture map data in accordance with claim 15 or claim 16;

determining the projection of a three dimensional model defined by said model data as perceived from said selected viewpoint; and

10 texture rendering said projection of said model utilising data said texture map data.

18. A method in accordance with claim 17 further comprising a step of outputting said generated image data of a subject from a selected viewpoint.

19. Apparatus for generating frequency component image data representative of image information of a plurality of different spacial frequencies of an image of at least part of a subject from a predefined view, the apparatus comprising:

15 means for receiving a plurality of images comprising image data representative of a subject from a plurality of view points;

20 means for determining for each of said plurality of received images, frequency component image data representative of image information of a plurality of different spacial frequencies within said images received by said receiving means and determining projections of said frequency component image data projected to said predefined view; and

25 means for utilising for each of said plurality of different spacial frequencies, said projections of frequency component image data to generate frequency component image data representative of spacial frequency information of an image of at least part of said subject from said predefined view.

20. Apparatus in accordance with claim 19, wherein said means for determining is arranged to determine frequency component image data representative of a lower spacial frequency by performing a blurring and sub-sampling operation utilizing said image data received by said receiving means.

21. Apparatus in accordance with claim 20 wherein said means for determining is arranged to determine frequency component image data representative of a higher frequency by performing a differencing operation utilising said lower frequency component image data and said image data received by said receiving means.

22. Apparatus in accordance with any of claims 19 to 21 wherein said means for determining frequency component data comprises:

30 means for determining from said plurality of received images the relative positions of the viewpoints of the subject represented by said images;

35 modeling means for determining from said plurality of images and said determined positions, model data representative of the surface of said subject; and

40 means for determining said projections utilising said position data and said model data of the surface of said subject.

23. Apparatus in accordance with claim 22, wherein said means for determining relative positions is arranged to determine position data indicative of the relative position of said plurality of viewpoints for said received images by matching features within said images to determine relative position and orientation of the viewpoints.

45 24. Apparatus in accordance with claims 22 or 23, wherein said modeling means is arranged to generate model data by:

determining the portions of an image corresponding to the subject; and

50 generating model data utilising a determination of the perimeters of said portion of said image and position data indicative of the viewpoint for said image.

25. Apparatus in accordance with any of claims 19 to 24 wherein said means for utilising is arranged to generate frequency component image data for different spatial frequencies by utilising each portion of said projections of frequency component image data for a frequency to determine frequency component image data representative of spatial frequency information of said spacial frequency for corresponding portions of said image.

55 26. Apparatus in accordance with claim 25, wherein said means for utilising is arranged to generate frequency

**EP 1 204 073 A2**

component image data for different spatial frequencies for an image from said corresponding portions of projections of frequency component image data in different ways for different spatial frequencies.

27. Apparatus in accordance with claims 25 and 26, wherein said means for utilising further comprises:

5 means for determining visibility data indicative the relative visibility of portions of the surface of said subject from said viewpoints;

10 means for associating portions of said projections of frequency component image data, with said visibility data; and

15 processing means for utilising said visibility data associated with portions of said projections to determine frequency component image data representative of frequency image information of an image of said subject from said predefined views.

20 28. Apparatus in accordance with claim 27, wherein said processing means is arranged to generate frequency component image data by calculating for each portion of said image data for a said spacial frequency a weighted average of corresponding portions of projections of frequency component image data for said portion of said image, weighted by said visibility data associated with said portions of image data.

25 29. Apparatus in accordance with claim 27, wherein said processing means is arranged to generate frequency component image data by selecting for each pixel of said image data for a said spacial frequency said corresponding portions of projections of frequency component image data associated with visibility data indicative of the highest visibility.

30 30. Apparatus for generating image data representative of an image of a subject from a predefined viewpoint, the apparatus comprising:

means for generating frequency component image data representative of image information of a plurality of different spacial frequencies of an image of a subject from said predefined view comprising apparatus in accordance with any of claims 19-29; and

35 means for generating image data representative said image of said subject utilising said frequency component image data.

30 31. Apparatus in accordance with claim 30, wherein said means for generating image data is arranged to determine for each portion of said image a determined the sum of frequency component image data for said different spacial frequencies.

35 32. Apparatus in accordance with claims 30 and 31 further comprising output means for outputting said generated image data.

40 33. Apparatus for generating image data of a subject from selected viewpoint comprising:

map generating means for generating texture map data for texture rendering a model of said subject, comprising apparatus in accordance with any of claims 30, 31 or 32;

means for determining the projection of three dimensional computer model as perceived from said selected viewpoint; and

45 means for texture rendering said projection utilising texture map data generated by said map generation means.

34. An apparatus for generating texture render data comprising:

45 a confidence image generator operable to generate a plurality of confidence images, said confidence images being representative of the relative visibility of parts of the surface of an object from defined viewpoints;

50 an association unit operable to associate each of said confidence images with a source image, at least one of the source images comprising an image other than an image of said object from the defined viewpoint represented by said confidence image; and

55 a texture generator operable to generate texture render data by selecting portions of said source images on the basis of corresponding portions of said associated confidence images.

35. A method of generating texture render data comprising the steps of:

55 generating a plurality of confidence images, said confidence images being representative of the relative visibility of parts of the surface of an object from defined viewpoints;

associating each of said confidence images with a source image, at least one of the source images

**EP 1 204 073 A2**

comprising an image other than an image of said object from the defined viewpoint represented by said confidence image; and

generating texture render data by selecting portions of said source images on the basis of corresponding portions of said associated confidence images.

- 5      36. A recording medium for storing computer implementable process steps for generating within a programmable computer, apparatus in accordance with any of claims 19 to 34.
37. A recording medium for storing computer implementable process steps for causing a programmable computer to perform a method in accordance with any claims 1 to 18 or 35.
- 10     38. A recording medium in accordance with claim 36 or claim 37 comprising a computer disc.
39. A computer disc in accordance with claim 38 comprising an optical, magneto-optical or magnetic disc.
40. A recording medium in accordance with claim 36 or claim 37, comprising an electrical signal transferred via the Internet.

15

20

25

30

35

40

45

50

55

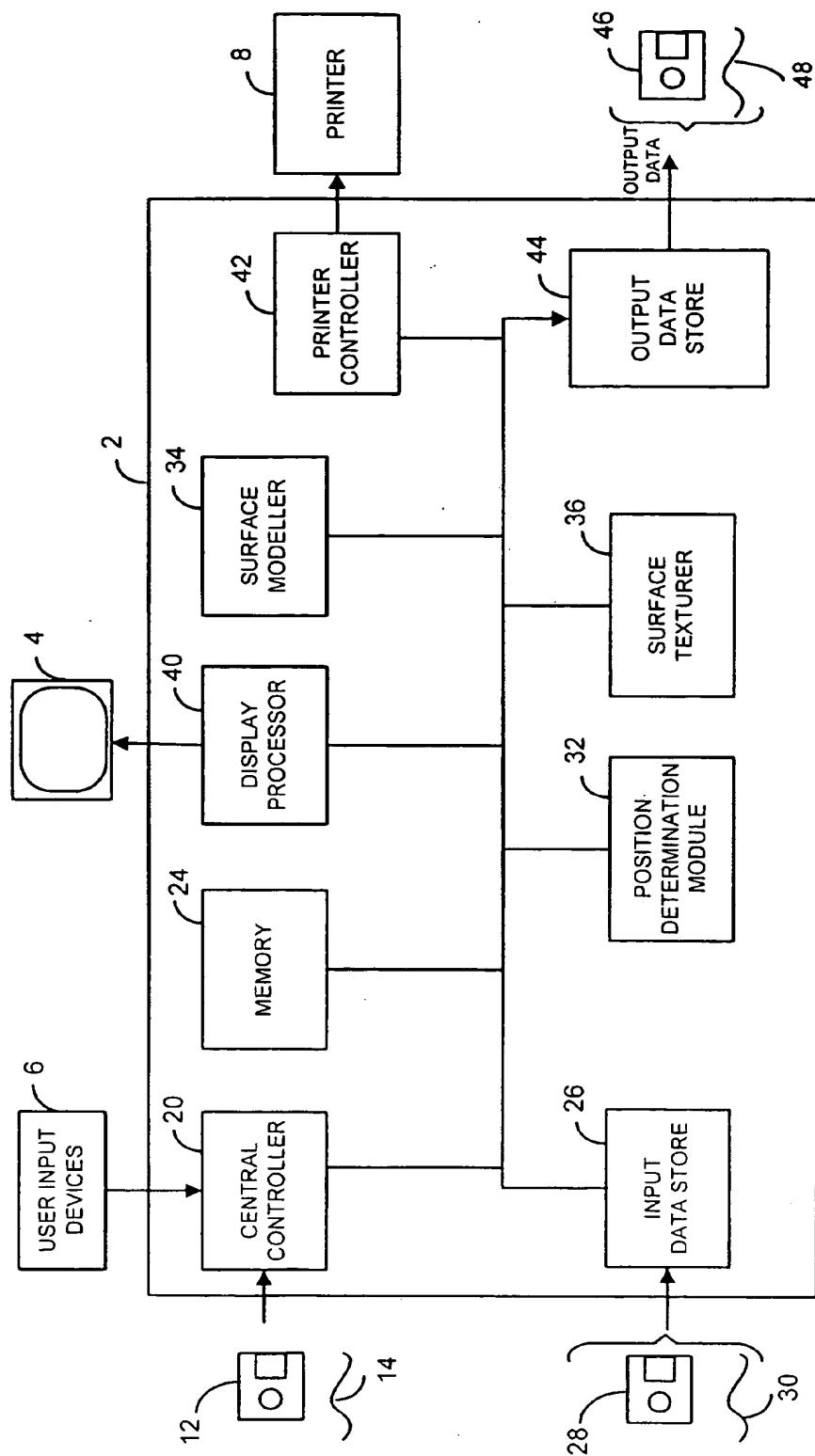
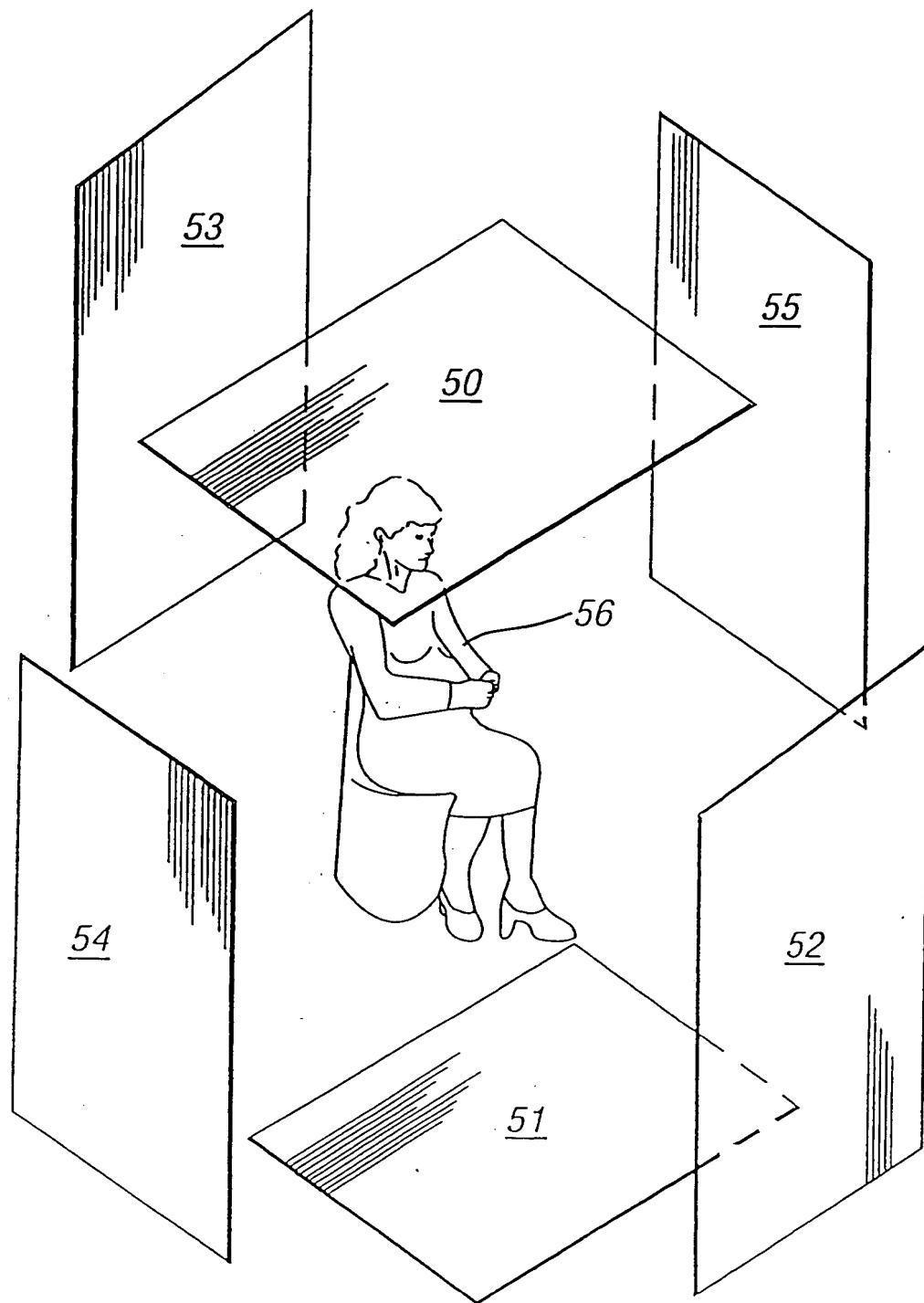


FIG. 1

FIG. 2



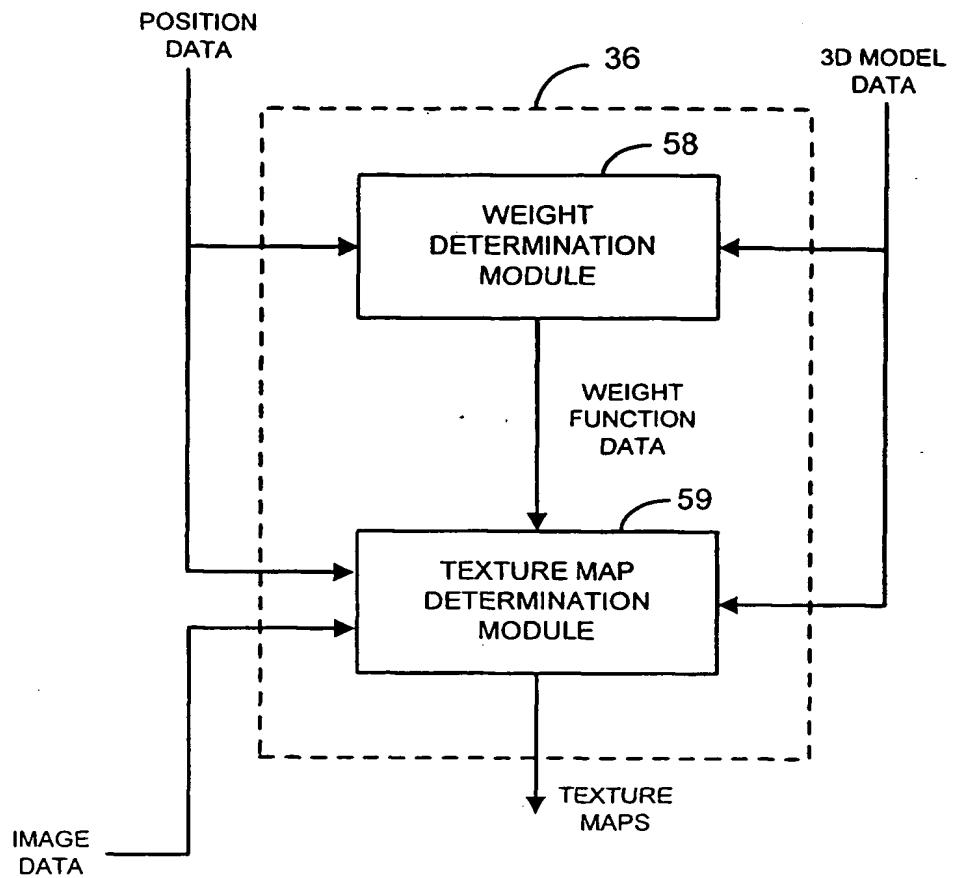


FIG. 3

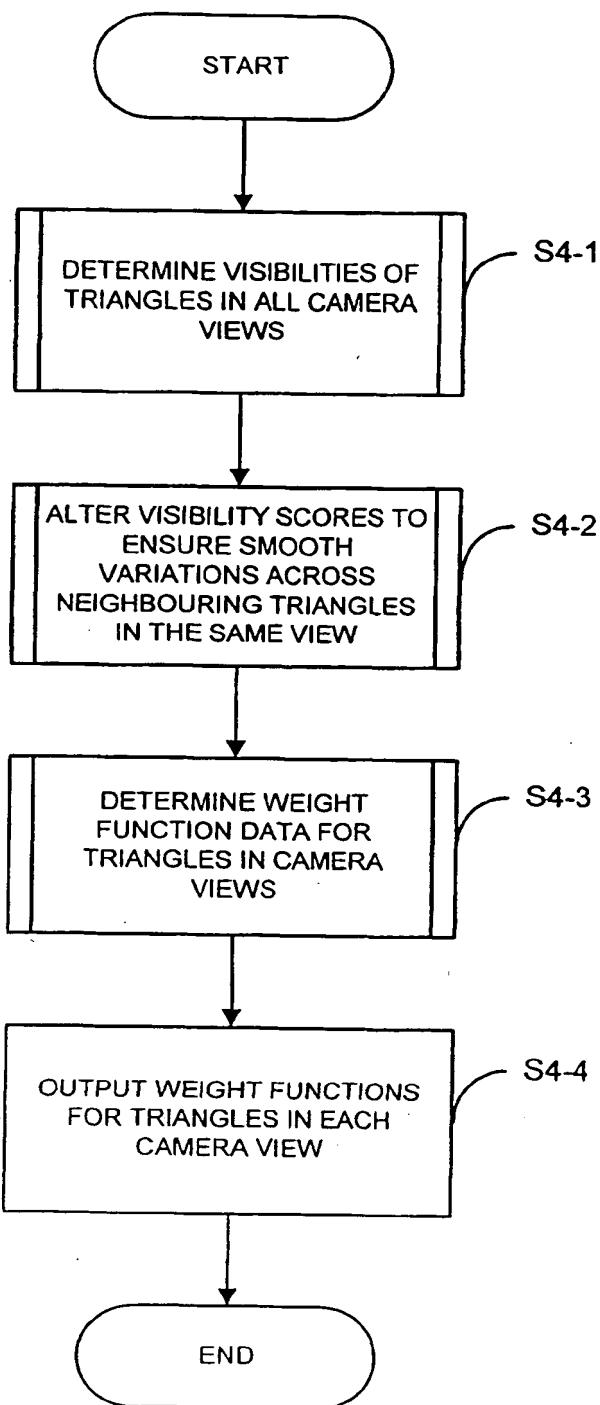


FIG. 4

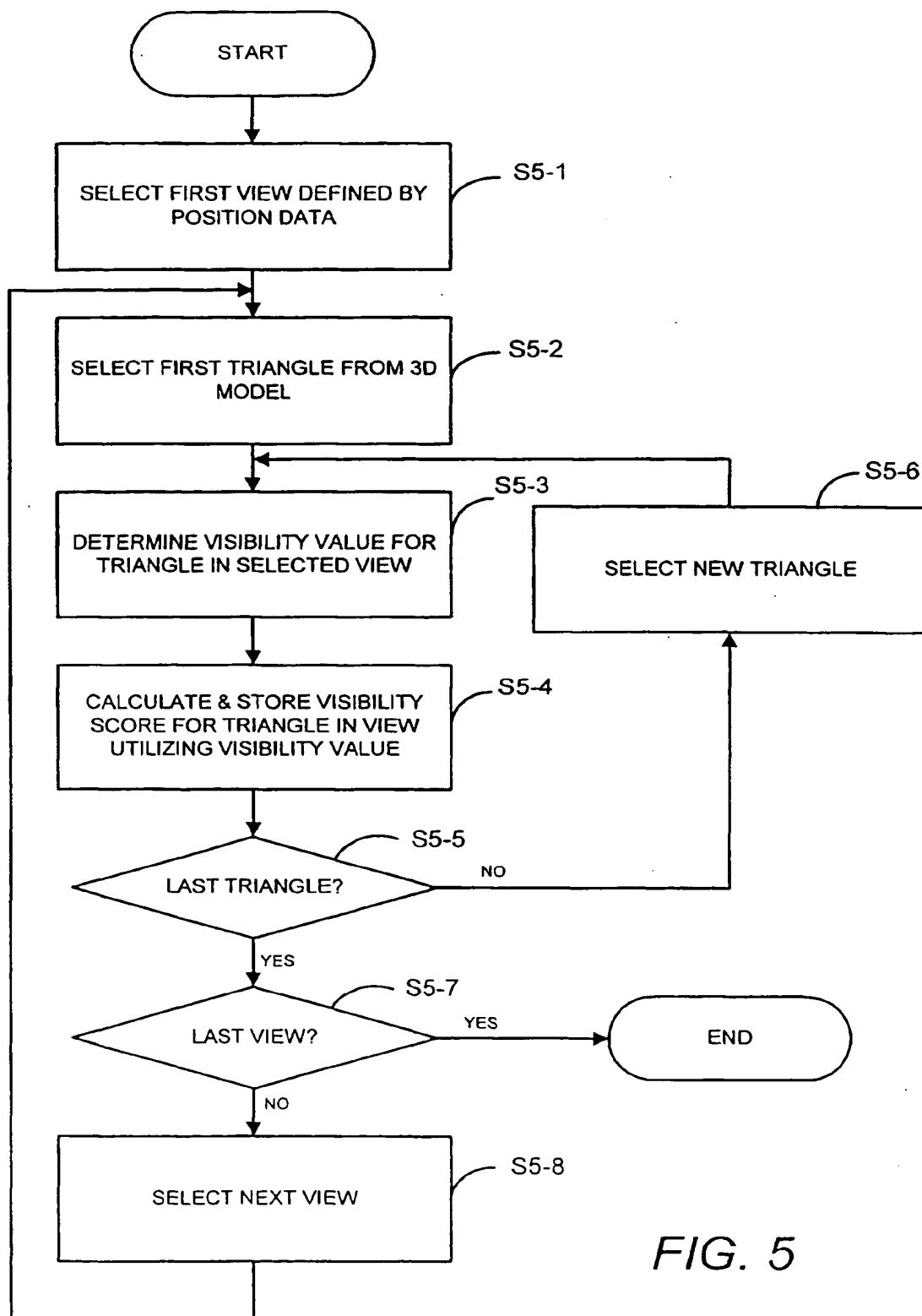


FIG. 5

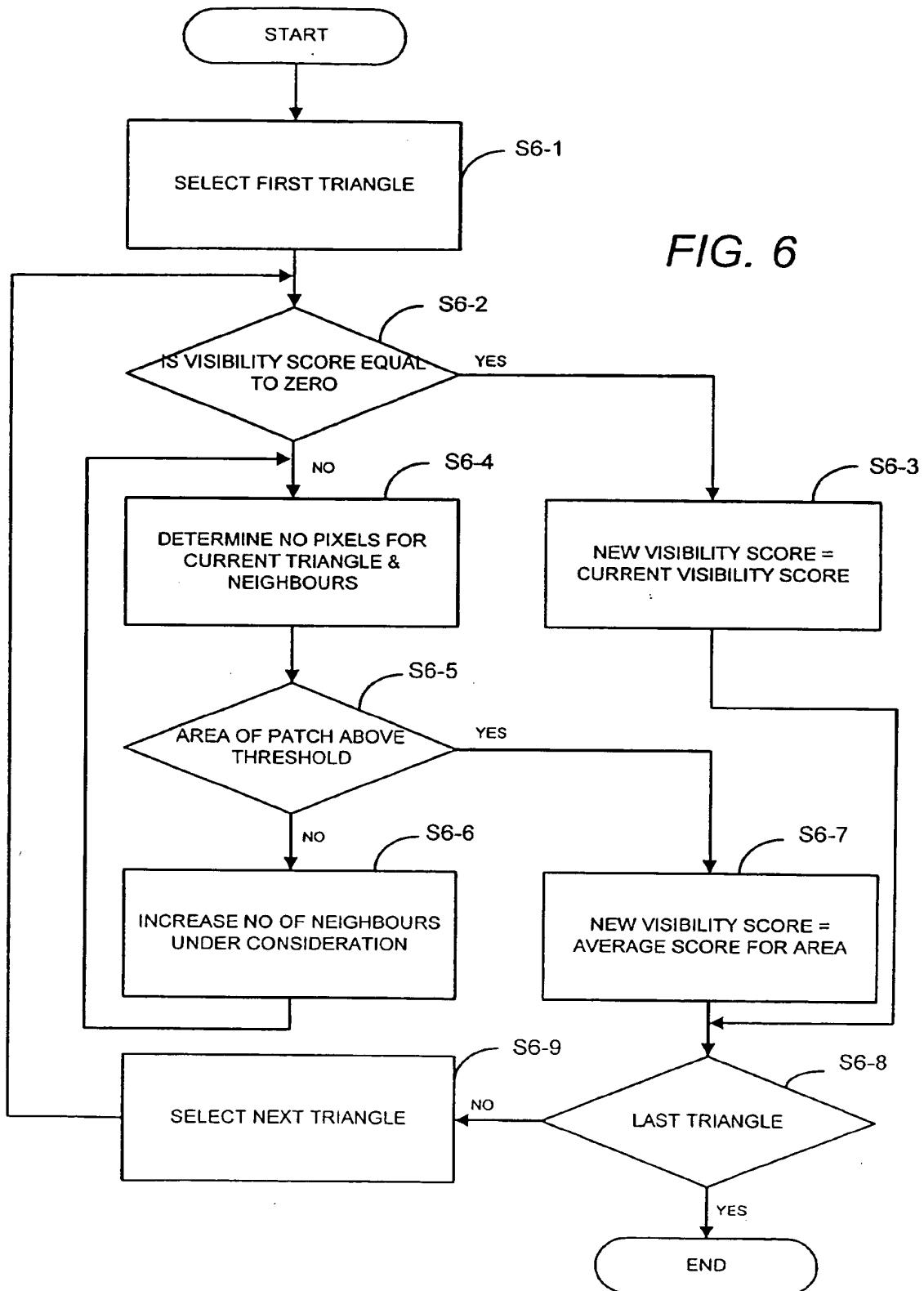


FIG. 6

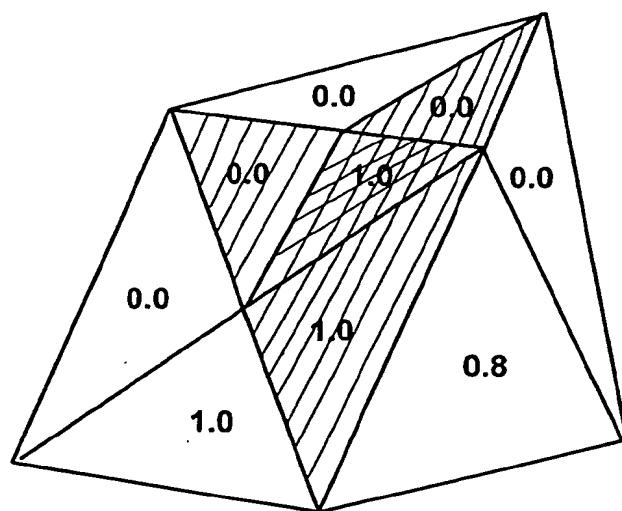


FIG. 7

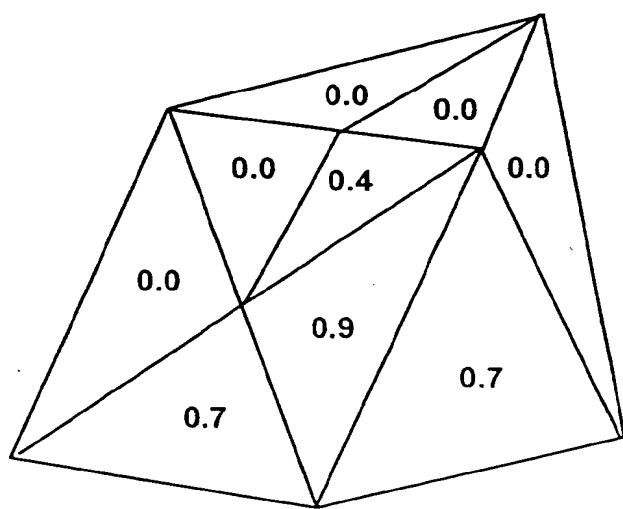


FIG. 8

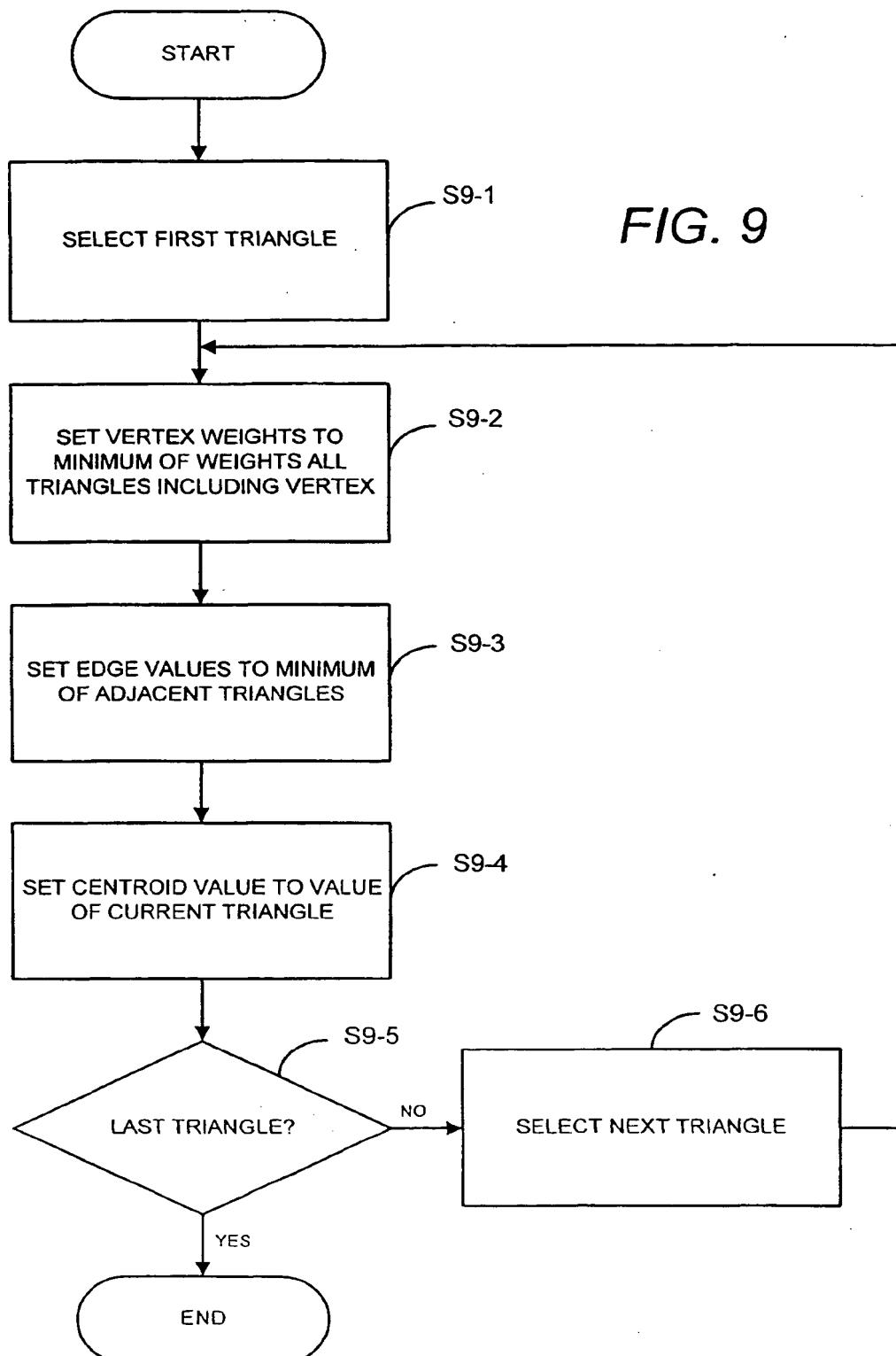


FIG. 9

EP 1 204 073 A2

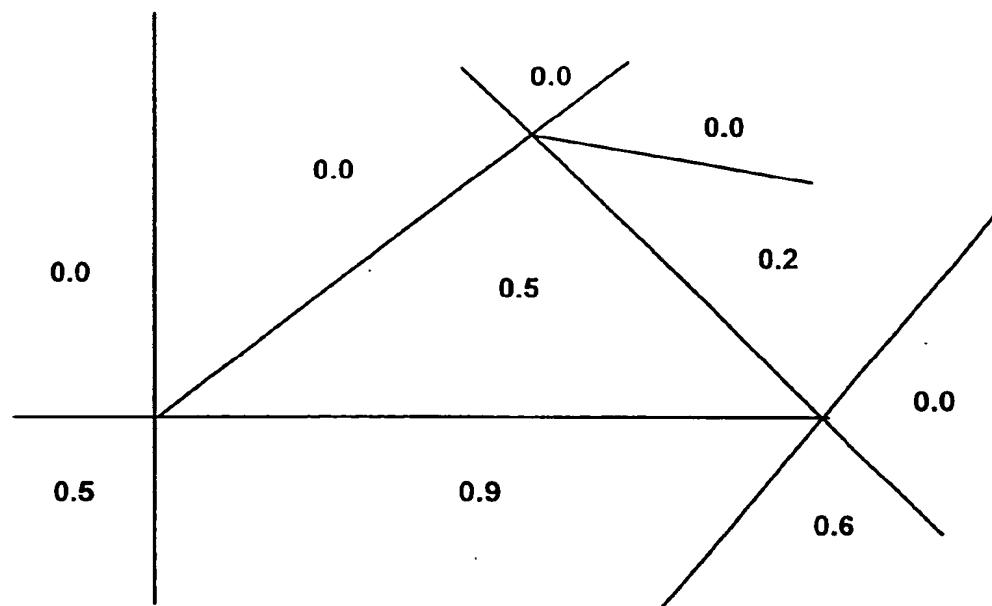


FIG. 10A

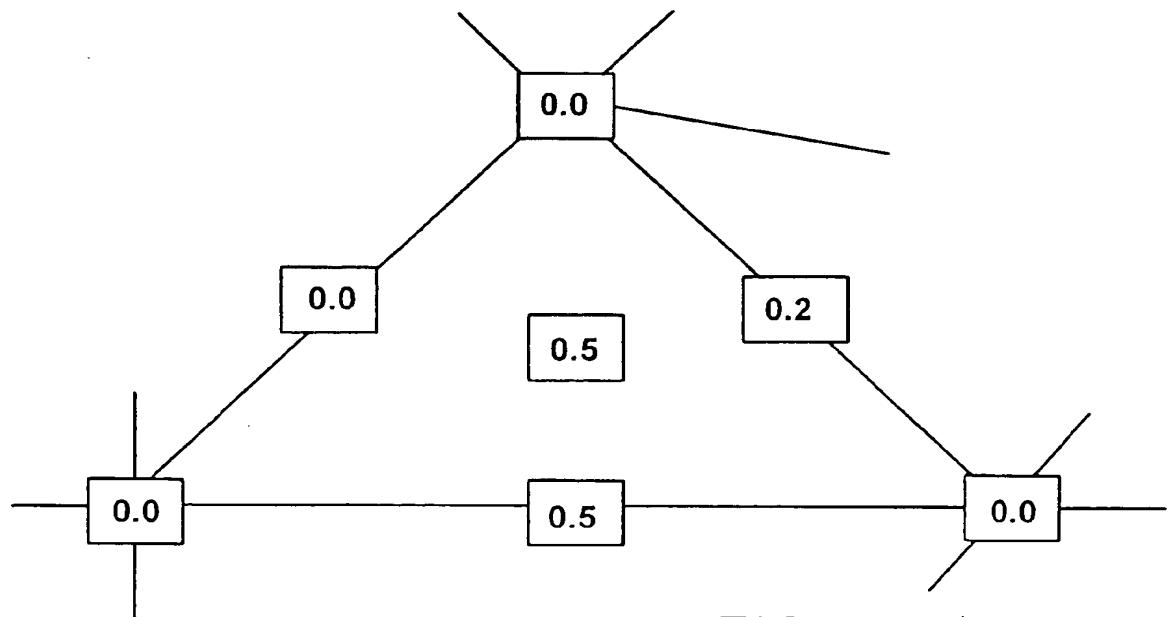


FIG. 10B

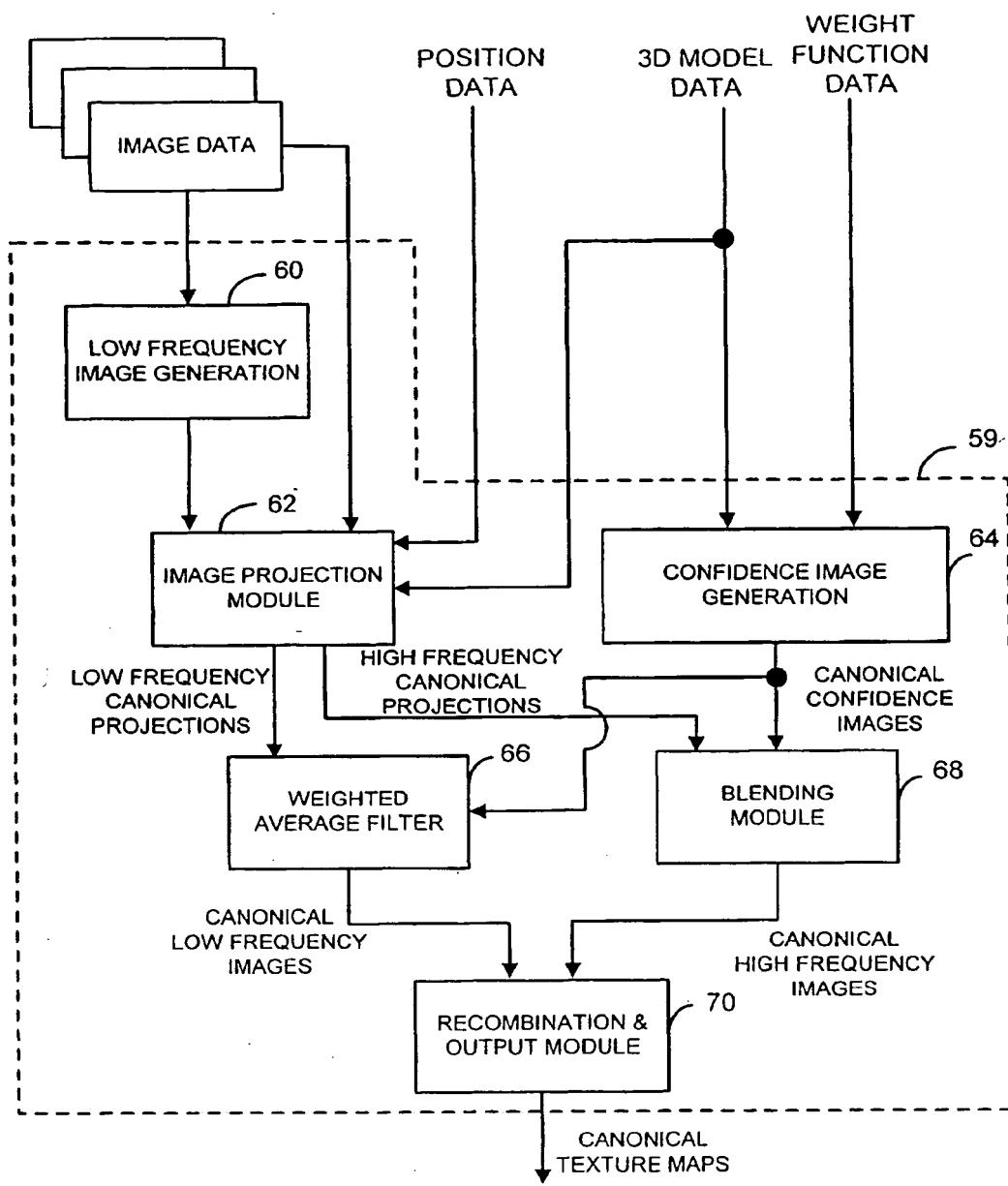


FIG. 11

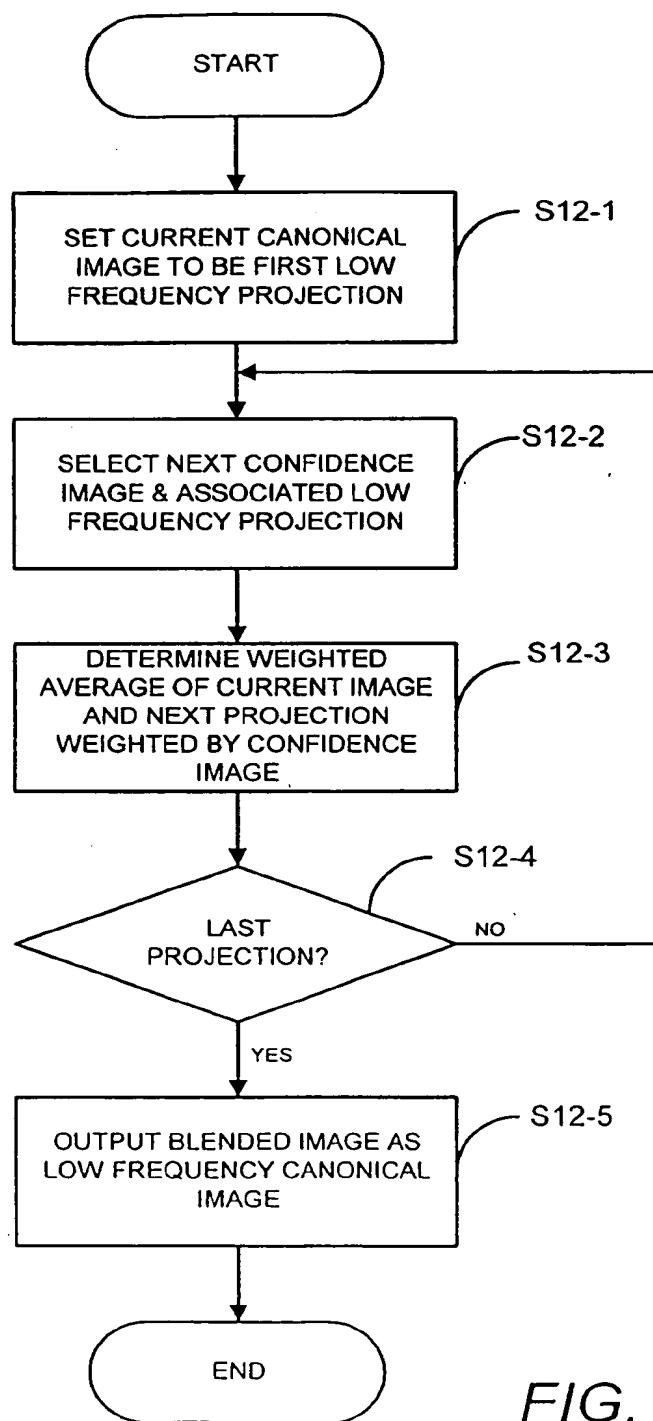


FIG. 12

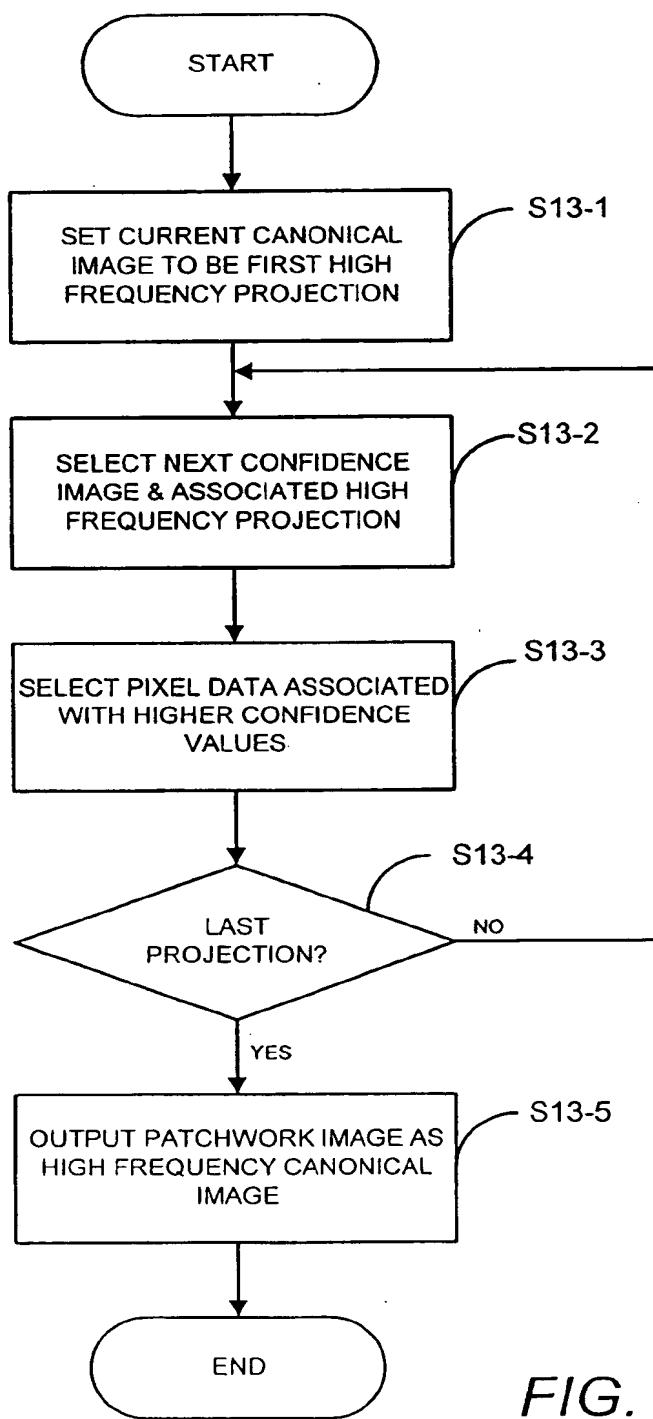


FIG. 13

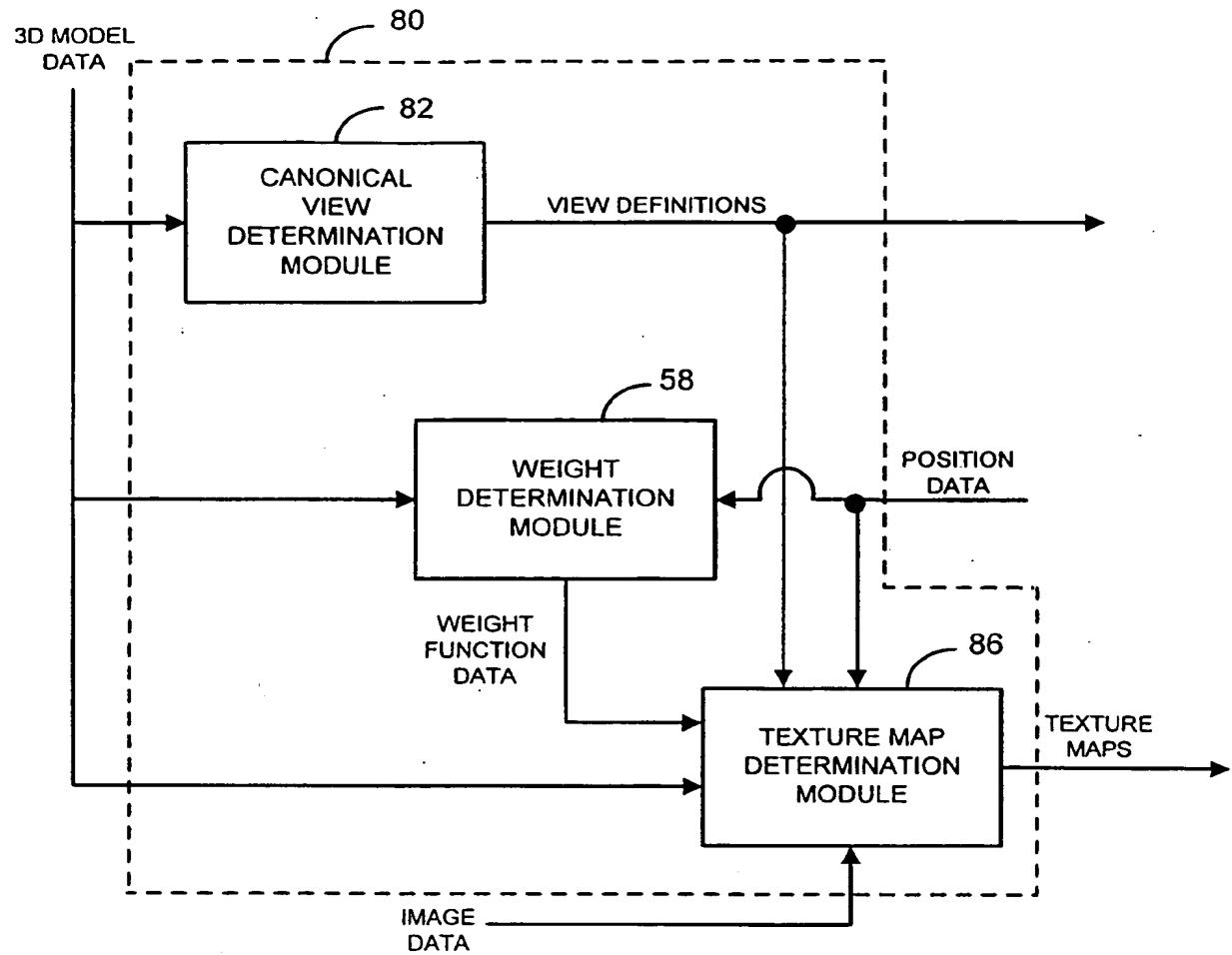


FIG. 14

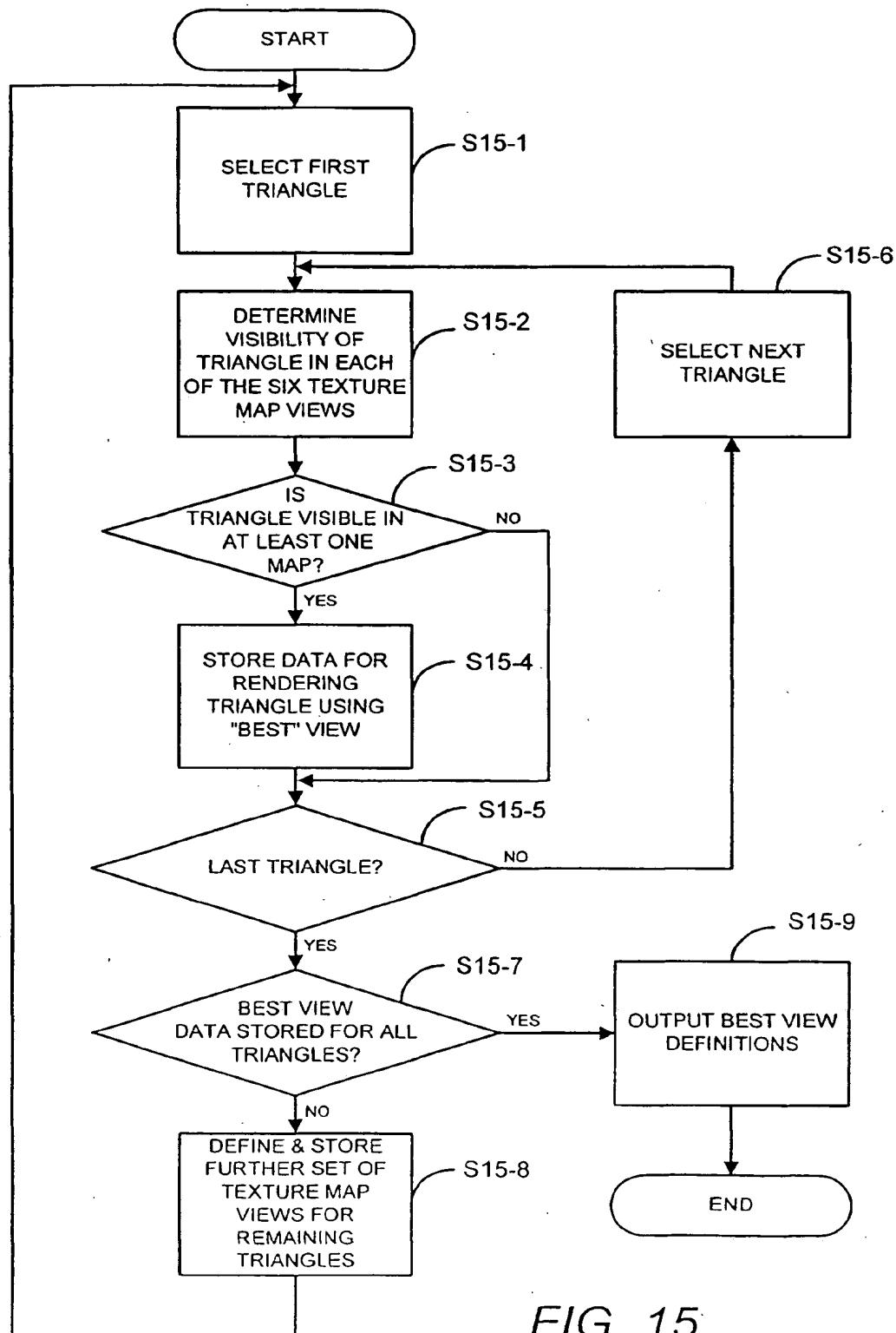


FIG. 15

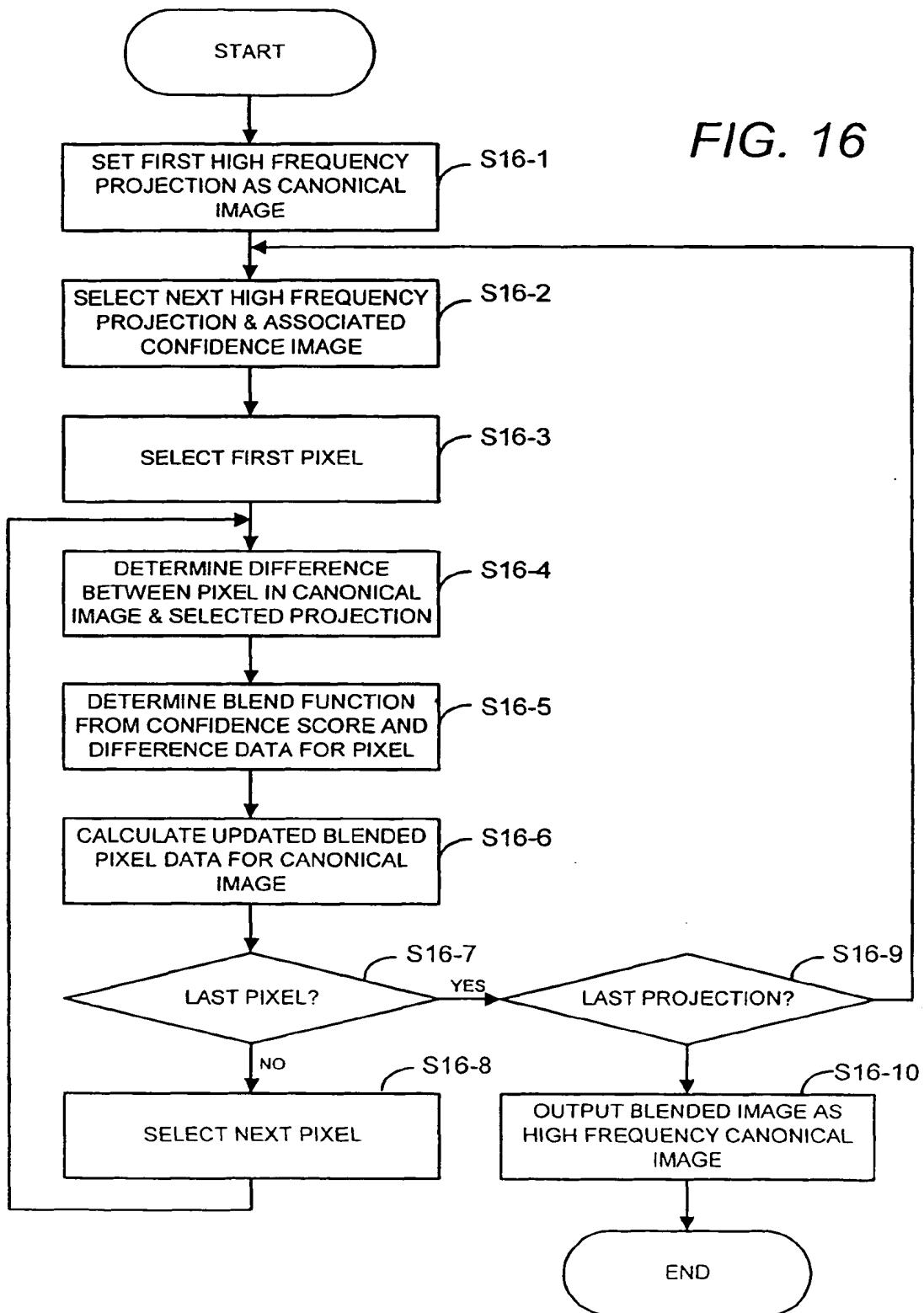


FIG. 16

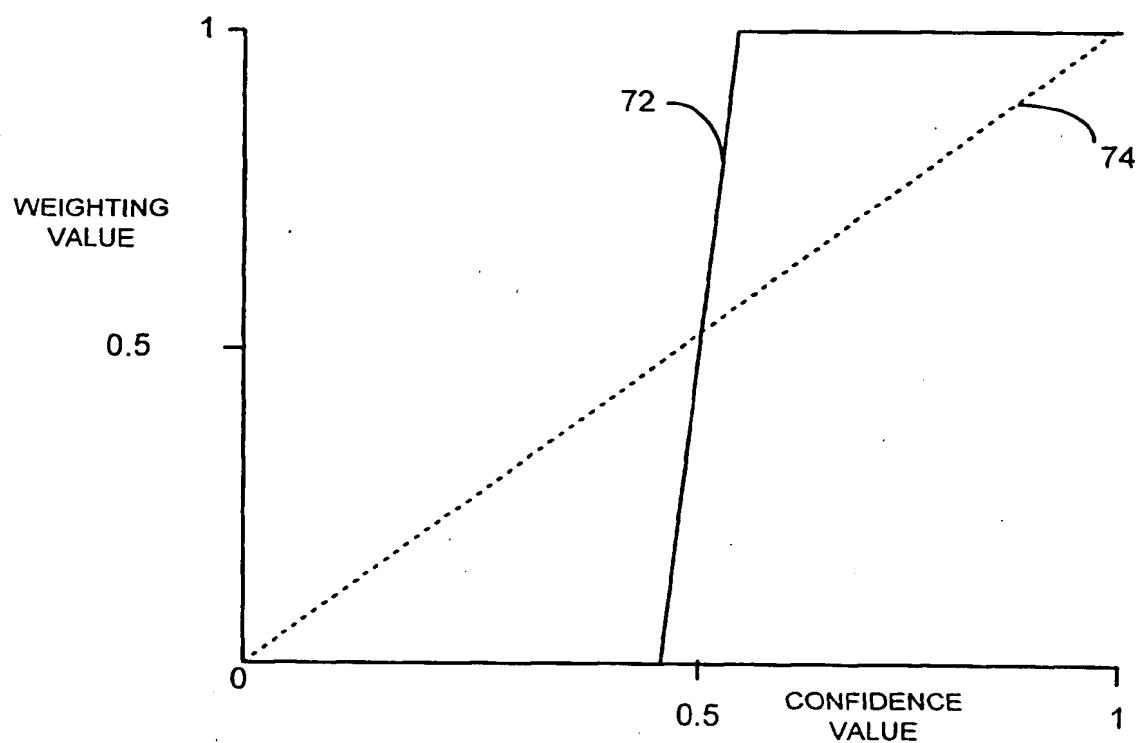


FIG. 17